SPECIFICITY OF NEUROFEEDBACK TRAINING IN ATTENTION-DEFICIT/HYPERACTIVITY DISORDER (ADHD) – MULTILEVEL MODELING OF EEG-LEARNING

Thesis (Cumulative Thesis)

Presented to the Faculty of Arts and Social Sciences of the University of Zurich for the Degree of Doctor of Philosophy

by Agnieszka Zuberer

Accepted in the spring semester 2017

on the recommendation of the doctoral committee:

Prof. Dr. Lutz Jäncke (main supervisor)

PD. Dr. Renate Drechsler

Zürich, 2017

Acknowledgements

I would like to thank all of those who helped me on my academic career. My greatest thanks go to PD. Dr. Renate Drechsler for her patience in teaching me over the past three years and supporting my research. Especially I would like to thank Prof. Dr. James Swanson for his fruitful stimulations on statistics. I owe Prof. Dr. Jäncke my greatest thanks for the inspirational stimulations for Neuroscience, especially on topics like attention and altered states of consciousness. I also would like to thank Franziska Minder, for always creating such a wonderful working atmosphere and being a friend in need.

I wish to express my greatest thanks to my family and to Charley M. Wu for constantly motivating me to learn more about things in the world.

TABLE OF CONTENTS

Summary

Although neurofeedback (NF) is a very prominent clinical intervention for attention-deficit/hyperactivity disorder (ADHD), the question persists if and how clinical improvement results from the clinical intervention. Many studies report an improvement of ADHD severity after treatment with Neurofeedback (NF); however, the specific effects of the treatment are barely examined. Disentangling specific from non-specific effects of a treatment is necessary to decide if an intervention is appropriate for the individual or/and how the intervention can be improved to achieve better treatment results. The aim of this thesis is to help elucidate these issues by first reviewing the literature on studies analyzing training performance and its relation to clinical outcome (paper 1, review). In the first experimental paper (paper 2), study results of the NF training are reported with a focus on the learning performance across multiple training sessions and its association to subject related (e.g., IQ, age) and treatment related (e.g., school versus clinical setting) factors. A mixed-effects modeling approach showed NF learning to be dependent on IQ, age and intake of stimulants. The aim of paper 3 was to examine how individual learning in NF is related to clinical outcome. By employing a mixed models approach, no association between NF learning and clinical outcome was found. In contrast, alpha power increments across training sessions were partially associated with clinical improvements.

Zusammenfassung

Obwohl Neurofeedback (NF) eine sehr bedeutende klinische Intervention für eine Aufmerksamkeits-Defizit-Hyperaktivitäts-Störung (ADHS) ist, bleibt die Frage bestehen, ob und wie klinische Verbesserungen aus der klinischen Intervention resultieren. Viele Studien berichten eine Verbesserung der ADHS Symptome nach einer Neurofeedback (NF) Behandlung; jedoch sind die spezifischen Effekte der Behandlung kaum untersucht. Das Herausarbeiten spezifischer von unspezifischer Behandlungseffekte ist notwendig, um zu entscheiden, ob eine Intervention für ein Individuum geeignet ist oder/und wie diese Intervention verbessert werden kann, um bessere Behandlungsresultate zu erzielen. Ziel dieser Arbeit ist es, diese Fragestellungen zu eruieren, indem zuerst die Literatur zu Studien, welche die Trainingsleistung und ihre Beziehung zu klinischen Werten berichten, aufgearbeitet wird (wissenschaftlicher Artikel Nr. 1). In der zweiten experimentellen Arbeit (wissenschaftlicher Artikel Nr. 2) werden Studien im Bezug auf das NF Lernen zusammengefasst, mit einem Fokus auf die Lernleistung über mehrere Sitzungen hinweg und seine Assoziation zu personenbezogenen (z.B. IQ, Alter) und behandlungsbezogenen (z.B. Schul-versus Klinikkontext) Faktoren. Analysen mit einem gemischten Modell Ansatz ergaben, dass NF Lernen von IQ, Alter und Einnahme von Stimulanzien abhängig ist. Das Ziel der dritten Arbeit war es zu untersuchen, wie das individuelle Lernen in NF mit den klinischen Werten in Bezug steht. Analysen mit gemischten Modellen ergaben keine Assoziation zwischen NF Lernen und klinischen Werten. Im Gegensatz dazu waren Alpha-Power Zunahmen über Trainingssitzungen teilweise mit klinischen Werten assoziiert.

1 Introduction

1.1 General introduction

Over the past years neurofeedback (NF) has gained increasing popularity as a training method for children and adults with ADHD. Analyses of NF training efficacy usually focus on clinical pre-post improvements, but ignore specificity, e.g. whether or not children with ADHD gain control over their brain activity during the training sessions. It is still unknown why some children are “good learners”, while others seem unable to regulate their brain activity. The goal of the present studies was to evaluate factors that might influence EEG-learning performance and to analyze the association between EEG-learning and clinical outcome.

1.2 Attention-Deficit/Hyperactivity Disorder

Attention deficit hyperactivity disorder (ADHD) is a childhood onset neurodevelopmental disorder, characterized by increased inattention, hyperactivity, and impulsivity as compared to typical developing children. Problems are present for more than six months, and cause problems in at least two settings (such as school, home, or recreational activities). Worldwide prevalence is about 5% of school-aged children with boys being overrepresented (Fayyad et al., 2007; Kessler, Adler, & al., 2006; Polanczyk, De Lima, Horta, Biederman, & Rohde, 2007) with relatively high heritability (Faraone, Biederman, & Mick, 2006). Often ADHD has comorbidities such as dyslexia and dyscalculia, conduct oppositional defiant and internalizing disorder (Willcutt et al., 2013; Willcutt et al., 1999). Symptoms are present for more than six months, and cause problems in at least two settings (such as school or home).

Research suggests treatment with stimulants to be effective for up to 14 months, whereas long term effectiveness is yet unclear (Huang, Tsai, & Guilleminault, 2011).

According to the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) by the American Psychiatric Association (APA, 1994), the diagnostic criteria include a) symptom thresholds, b) criteria for pervasiveness, c) persistence of impairment, and d) age of onset. Whereas DSM-IV criteria have a clear distinction between three ADHD subtypes, the subtype distinction and age of onset criteria of the DSM-5 are less strict. For the present study DSM-IV criteria were used, because at study onset the employed rating scales were not yet adjusted for the DSM-5. Subtypes according to DSM-IV criteria comprise I. the predominantly hyperactive-impulsive, II. the predominantly inattentive, and III. the combined subtype. This subdivision has been criticized to be biased because diagnostic items do not take into account developmental aspects. Furthermore, diverse informants may provide differing information on predominant symptoms. A research diagnosis entails beside symptom ratings (from parents and teachers) also neuropsychological tests, although these are not required for clinical diagnosis.

1.3 Neurofeedback

Neurofeedback is a form of brain-computer interface, employed to promote the perception and control of specific aspects of brain activity by providing real-time feedback of a specific EEG-parameter. The learning process is based on an operant conditioning paradigm. Neurofeedback has been employed since the 70ies (Lubar & Shouse, 1976), but controlled studies on neurofeedback in ADHD exist since around 20 years (Lubar et al., 1995). In ADHD, the two most prominent types of NF are frequency-band NF (FR-NF) and NF of Slow Cortical Potentials (SCPs). In FR-NF a

specific EEG frequency band or ratio of frequency bands is fed back to the subject. This type of NF aims at tonic aspects of the EEG aiming to achieve changes over a relatively long time period (30 -60 minutes). The most prominent FR- NF protocols aim to reduce Theta/Beta ratio or either increase Beta or reduce Theta (see Strehl et al., 2013).

1.1. Neurofeedback of Slow Cortical Potentials

Whereas in FR-NF the subject tries to reach one specific state, in SCP-NF the aim is to switch between a state of activation / alertness and a state of deactivation / relaxation within shorter time periods of around 10 seconds. SCPs are associated with very low frequency range of < 0.1 Hz and are very slow EEG fluctuations (Monto, Palva, Voipio, & Palva, 2008). They originate in the apical dendritic layers of the neocortex and reflect synchronized depolarization of large groups of neuronal assemblies (Birbaumer & Elbert, 1990). According to Birbaumer's threshold regulation model of cortical excitation (1990), negative and positive SCPs are associated with an activated or deactivated state, respectively.

2 Methods

2.1 How to measure training efficacy and specificity?

The employment of a proper, randomized control group, is one of the most crucial aspects for measuring treatment specificity. However, the choice of the type of control group constitutes a trade-off between factors that are controlled of, and factors that are not. To start with, semi-active control conditions are mainly treatment groups where unspecific effects such as trainer-patient interactions are controlled for. Such conditions could be for instance a training group, where the amount of training sessions and training frequency is comparable to the treatment condition, whilst the parameter to be trained groups differ (e.g. muscle vs. brain activity; see Maurizio et al., 2013). For instance, one would expect muscle training to have a weaker impact on ADHD symptom ratings than a neurofeedback training which directly aims at the improvement of attention and self-regulation.

In contrast, an active control condition would be a treatment group, where no a priori assumptions about the preference of one treatment over the other can be made (e.g. when contrasting neurofeedback with a cognitive training group).

At last, another approach is to employ a placebo group, where, in case of sham neurofeedback, the parameter to be fed back to the subject is set random and unrelated to brain activity. Expectancy effects due to having electrodes attached to the head would be controlled for in this situation (see Arns et al., 2014 for discussion on this matter). However, there are ethical objections to the treatment, namely that subjects would invest time in a treatment that probably would not be effective. Already the expectancy, to be in the possible placebo group might bias the results, and finally,

the random feedback might also in some cases coincidently go along the actual brain activity, which would contradict to the principles of a placebo group.

Another way to examine training specificity is to relate clinical outcomes to the learning process in the treatment group. One would expect that if the training is the cause for clinical improvement, a poor learner would show no or only weak improvements in the clinical domain. This approach does not include another control group.

2.2 Multilevel Modeling

Over the past years, there has been a critical discussion on p-values as the “gold standard” for statistical validity, probably being one major contribution to why results of many psychological studies could not be reproduced (Greenland et al., 2016; Dahiru et al., 2008). As a result, one of the biggest formations of researchers was founded to reproduce previous findings in psychological and to derive new scientific guidelines that address both researchers and editors (Jarrett, 2015; Weir, 2016).

Most basic statistical methods (such as multivariate variance analysis) assume that observations are randomly and independently sampled from a population. However, the independence assumption is often violated. For instance, violation can occur when data structures are nested, such as children are nested within classrooms, which in turn are nested within schools; these schools in return would be nested in districts with a specific socio-economical background and so on. Independence can also easily be violated when analyzing longitudinal data with multiple measurements of a subject across time. If independence gets violated standard errors of parameter estimates are overestimated, p-values inflated, and the chance to detect significant effects decreases.

One way to take into account these methodological shortcomings is to employ multilevel modeling approaches by investigating covariates of units of different levels (see Figure 1).

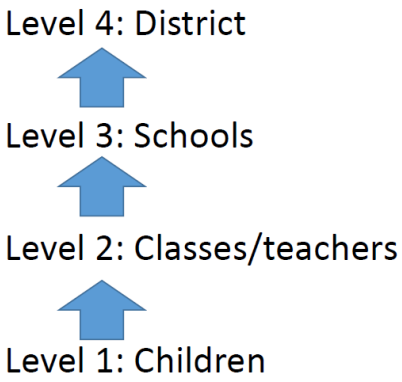

Figure 1: Visualization of hierarchical structure of effects in multi-level modeling.

Linear mixed effects are an extended form of linear regression analysis, where the variability between and across subjects is taken into account. The common equation for a linear regression does not distinguish between subgroup or individual effects and attributes this inter-subject variability and other unexplainable effects to a single error term ε (see equation 1). In the random-coefficient model this error term gets extended by either a term for a random intercept or a term for random slope for each individual / subgroup or both (random intercept and random slope).

$$v_{ij} = \gamma_{00} \cdot \gamma_{10} \cdot X_{1,ij} + b_{0j} + b_{1j} \cdot X_{i,ij} + \epsilon_{ij} \qquad (1)$$

Equation 1: Random coefficient regression model equation for individual I (i = 1 .. n_j) in group j (j = 1…N)

Y_{00} overall intercept

γ_{10} is the overall effect of x_1 on y.
b_{0j} is the random deviation of each group from the overall intercept
b_{1j} is the random deviation of each group from the overall slope
ε error term

Thus, for measuring treatment efficacy, mixed models allow not only to examine treatments effects on a group level (treatment versus control group), but also to control for the subject intercept/offset (clinical severity before treatment varies between subjects / groups) and/or the variability of the slope: degrees of change of clinical severity over time varies (Bates et al., 2003).

2.3 Theoretical Background and Research Questions

Although NF has been employed for over two decades in controlled studies in ADHD, there is still no consensus, if treatment related clinical improvements actually originate from the treatment itself or from unspecific factors (such as trainer-subject relationship, placebo-effects). This problem is mainly due to the lack of controlled studies measuring if and how the actual NF-learning is related to clinical improvements (see Paper 1). For instance, if a subject shows clinical improvement but proofs to be a poor NF-learner, this would give reason to believe that the clinical improvement probably is not related to the treatment itself. From here on, the ability to modulate the NF-parameter will be called “NF learning” without presumptions about learning success, in line e.g. with Nan et al. (2015) and Zambotti et al. (2012). Over the past decades of NF-research in ADHD, only around 20 studies included a measure for NF-learning and they employed very different methods. That exacerbates conclusions on a reliable learner rate in NF. Even less studies related NF-learning to the treatment related change in clinical severity (see Paper 1). Thus, the main purpose of this thesis was to develop a more

reliable approach to measure treatment efficacy which resulted in following publication topics:

Paper 1 is a systematic review on NF-treatment efficacy and specificity in ADHD. In this paper, studies were reviewed that examined either NF-learning alone or NF-learning in relation to changes in clinical severity. More specifically, the different units of measurements, different analytical approaches to determine both the degree of NF-learning and its classification into good and poor NF-learning were summarized.

Paper two deals with the analysis of predictors for NF-learning employing a methodological approach that also takes into account intra-and inter-subject variability (mixed effect modeling). By examining subject- and context- related variables (such as age, IQ, stimulants intake, clinic versus school setting) the aim was to gain a more reliable measure for NF-learning and its predictors.

In paper three, the learning performance that was determined in the latter (second) paper was employed to predict the change in clinical severity across three rating time points (three months before training onset, directly before training onset, directly after training end). In addition, the change in resting alpha power measured before and after each training session was used as a further predictor for change in clinical severity. [1]

3 Paper 1 - Are treatment effects of neurofeedback training in children with ADHD related to the successful regulation of brain activity? A review on the learning of regulation of brain activity and a contribution to the discussion on specificity

Published in *Frontiers in Human Neuroscience, 9*(March), 135.

Agnieszka Zuberer[1], Daniel Brandeis[1-4] & Renate Drechsler[1]

1 Department of Child and Adolescent Psychiatry, University of Zurich, Zurich, Switzerland
2 Neuroscience Center Zurich, University of Zurich and ETH Zurich, Switzerland
3 Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim/ Heidelberg University, Mannheim, Germany
4 Medical Faculty Mannheim/ Heidelberg University, Mannheim, Germany

3.1 Abstract

While issues of efficacy and specificity are crucial for the future of neurofeedback training, there may be alternative designs and control analyses to circumvent the methodological and ethical problems associated with double-blind placebo studies. Surprisingly, most NF studies do not report the most immediate result of their NF training, i.e. whether or not children with ADHD gain control over their brain activity during the training sessions. For the investigation of specificity, however, it seems essential to analyze the learning and adaptation processes that take place in the course of the training and to relate improvements in self-regulated brain activity across training sessions to behavioral, neuropsychological and electrophysiological

outcomes. To this aim, a review of studies on neurofeedback training with ADHD patients which include the analysis of learning across training sessions or relate training performance to outcome is presented. Methods on how to evaluate and quantify learning of EEG regulation over time are discussed. “Non-learning” has been reported in a small number of ADHD-studies, but has not been a focus of general methodological discussion so far. For this reason, selected results from the brain-computer interface (BCI) research on the so-called “brain-computer illiteracy”, the inability to gain control over one’s brain activity, are also included. It is concluded that in the discussion on specificity, more attention should be devoted to the analysis of EEG regulation performance in the course of the training and its impact on clinical outcome. It is necessary to improve the knowledge on characteristic cross-session and within-session learning trajectories in ADHD and to provide the best conditions for learning.

3.2 Introduction

Recent meta-analyses and reviews have evaluated the efficacy of neurofeedback training in children and have concluded that there is a need for more placebo-controlled studies in ADHD research with better blinding of raters and possibly also of trainers (Lofthouse, Arnold, Hersch, Hurt, & DeBeus, 2012; Sonuga-Barke et al., 2013). Placebo control, often used interchangeably with sham (Heywood & Beale, 2003; Van Dongen-Boomsma, Vollebregt, Slaats-Willemse, & Buitelaar, 2013) or mock (Tobias Egner, Strawson, & Gruzelier, 2002) feedback in this context, lacks only the active core component, namely the consistent feedback contingent upon specific EEG patterns, and appears indistinguishable from the neurofeedback condition. This typically implies that non-contingent sham feedback is provided to the participant during the training,

either by frequently changing contingencies with real data (e.g. Heywood & Beale, 2003), by using simulated EEG-like data or feedback (Logemann, Lansbergen, Van Os, Böcker, & Kenemans, 2010; Van Dongen-Boomsma et al., 2013), or pre-recorded data, which all may be combined with contingent feedback of real artefacts (Kerson, 2013). While placebo control and pre-post analyses of change on clinical, neuropsychological and electrophysiological levels would appear to be the first choice with regard to efficacy, it may be questioned whether they constitute the best method for investigating the specificity of NF. Although placebo control aims to control for all non-specific influences of the training setting, such as learning to sit still, improved personal well-being due to the positive relation to the therapist, or positive expectations, it entails methodological limitations. Sham feedback fails to control for generic and non-specific learning effects, i.e. by the experience of improvement and progressive mastery, of self-efficacy, and increase of control which may be induced by any kind of biofeedback. Although sham neurofeedback using slowly alternating contingencies with different frequencies may allow at least piecewise learning (Hoedlmoser et al., 2008), alternative placebo-type control conditions such as EMG biofeedback (Bakhshayesh, Hänsch, Wyschkon, Rezai, & Esser, 2011; Maurizio et al., 2014a), or feedback from a distinct control region as in neuroimaging (Caria et al., 2007) provide better control for progressive learning.

More importantly, with regard to specificity, neither placebo control nor any other type of control condition can provide positive proof that successful learning of EEG regulation in the active condition is responsible for clinical improvements. To that aim, it would be necessary to demonstrate that learning of EEG-regulation occurred during the training and that the NF-training success, in the sense of successfully learned self-regulation of brain activity across time, is related to positive outcome on the clinical,

neuropsychological or electrophysiological level (Holtmann, Sonuga-Barke, Cortese, & Brandeis, 2014). Adequate control for the generic effects of learning would then require successful learning at a similar rate in the control condition.

In addition, for the time being, the effects which might be induced by sham feedback remain poorly understood. This may be particularly relevant for individuals with ADHD, who according to the ADHD literature, may display problems with self-perception in various different ways: A sizable portion of children with ADHD show an inappropriate overestimation of self-efficacy and ability, the so-called illusory positive bias (Owens, Goldfine, Evangelista, Hoza, & Kaiser, 2007). Other studies have demonstrated feelings of low self-efficacy and low self-esteem in patients with ADHD (Mazzone et al., 2013; Newark & Stieglitz, 2010) which usually leads to a negative bias in self-perception. In addition, patients with ADHD seem to display problems with the self-perception of internal states (Donfrancesco et al., 2013). Many children with ADHD may be unaware of how it feels to be in an alert and focused state of mind. Thus, providing ADHD patients with sham feedback could prevent them from developing a more adequate self-perception or lead them to mistrust their intuition. Although the findings from sham neurofeedback control conditions suggest no detrimental effects regarding core ADHD symptoms, effects on self-perception remain to be tested directly. Also from this perspective, NF studies which use genuine neurofeedback and which examine whether learning of self-regulated EEG activity actually occurred during the training, may present a better alternative in order to investigate the specificity of NF than placebo controlled studies.

In this paper, we will present a short review of NF-studies with ADHD patients in which learning of EEG regulation was analyzed and we will discuss methods how to evaluate and quantify learning of EEG regulation over time. Among the many varieties of NF

protocols with ADHD (Hurt, Arnold, & Lofthouse, 2014), the training of frequency bands (NF-FB) and the training of slow cortical potentials (NF-SCP) are the best scientifically evaluated and will therefore be the focus of the following review (Table 1). We will additionally refer to studies with Q-EEG-training and with healthy participants or clinical groups other than ADHD in order to illustrate a respective method (Figure 1).

3.3 ADHD neurofeedback protocols and learning of EEG self-regulation

We identified 15 published NF group studies with ADHD children which include the analysis of EEG regulation learning across training sessions (Table 1). The majority of these studies used NF training of the frequency bands (NF-FR) and central electrodes. During NF-FR, subjects are provided with continuous (visual or/and audio) feedback and are positively reinforced as long as the spectral activity of the targeted EEG-frequency band or the ratio of specific frequency bands stays below (or above, respectively) a pre-defined threshold. As soon as the threshold is passed, the feedback stimulus changes, announcing that the subject has reached an undesired state. A classic ADHD study protocol aims to decrease theta activity and increase beta activity (Bakhshayesh et al., 2011; Leins et al., 2007b; Lubar, Swartwood, Swartwood, & O'Donnell, 1995). Another characteristic protocol for ADHD aims at increasing the sensorimotor rhythm (SMR) (Kropotov et al., 2005; Russell-Chapin et al., 2013), which is known to play an important role for motor excitability (Pfurtscheller, Stancák, & Neuper, 1996; Sterman, Howe, & Macdonald, 1970). While these frequency specific protocols are usually employed with the aim of obtaining "normalization" of characteristic spectral EEG abnormalities in ADHD, a more recent rationale is to train "regulation" of spectral EEG activity instead (Holtmann et al., 2009, 2014). This change in perspective is based on research that failed to find consistently abnormal or

characteristic EEG frequency patterns in children with ADHD at group level (Liechti et al., 2013). Consequently, some NF-FR protocols alternate between phases of up- and down-regulation which is consistent with the typical approach in SCP regulation (Leins et al., 2007b; Liechti et al., 2012; Maurizio et al., 2014b). In contrast, QEEG NF training (and / or z-score training) and other individualized NF protocols assume EEG abnormalities compared to normative data, which are trained in order to reach normalization (Hillard, El-Baz, Sears, Tasman, & Sokhadze, 2013a; Vollebregt, van Dongen-Boomsma, Buitelaar, & Slaats-Willemse, 2014).

In six out of 15 studies, NF of the slow cortical potentials (NF-SCP) was used (Drechsler et al., 2007; Gevensleben, Albrecht, et al., 2014; U. Strehl et al., 2006b), sometimes in combination (Liechti et al., 2012, Maurizio et al. 2014) or contrasted with NF-FR (Leins et al., 2007). SCPs are shifts in electro-cortical potentials which are thought to index the regulation of cortical excitability. NF-SCP trials are short, at about 8 sec., and participants are instructed to enhance activation (negativity trials) or reduce activation (positivity trials) relative to the baseline measured at the beginning of each trial. The magnitude of a produced negative amplitude reflects the amount of resources allocated to prepare a motor or cognitive response while a shift towards the positive polarity reflects a decrease in cortical excitability, which is in turn associated with a reduced responsiveness (Birbaumer & Elbert, 1990).

Table 1. ADHD Neurofeedback studies analyzing learning of EEG regulation

Study, NF-participants	**Protocol, electrode sites, no. of sessions,**	**Learning parameter / criterion for good performance**	**Learner rates/ learning outcome**	**Association between NF-learning and outcome gains**
Lubar et al., 1995 N = 17	Theta↓/ Beta↑ (bipolar electrodes situated halfway between Cz and Pz and halfway between Fz and Pz); 40 sessions	MP Theta/Beta / significant positive correlation between sessional learning parameter and session number	65 % learners	Stronger improvement in attentional test (TOVA) in learners than non-learners
Kropotov et al., 2005 N = 86	Beta↑ (C3-Fz); SMR↑ (C4-Pz); 15-22 sessions	At least 25 % increase of within sessional Beta- or SMR-power relative to resting-BL at the 1st session/ of at least 60 % of successful sessions	82.5 % learners	Improvements of ADHD symptoms and of Go/Nogo response-time and Go/Nogo SD
Strehl et al., 2006 N = 25 (Gani et al., 2008: 2-years-follow up)	SCP ↑↓ (Cz); 30 sessions (3 blocks of 10); follow-up sessions 31-33 (after 6 months)	MA of negativity trials during FB and TF, difference in MA between positivity and negativity trials / Good and poor learners based on median split of mean difference between MA of positivity and negativity trials at 3rd training phase	MA negativity trials: 2nd session < last session 2nd session < follow-up Difference between MA of positivity and negativity trials: at follow-up↑	Good TF-performance (difference between MA of positive and negative trials, sessions 21-30) is associated with clinical improvement only in good learners
Drechsler et al., 2007 N = 17 (Doehnert et al., 2008)	SCP ↑↓ (Cz); 30 sessions	MA of negativity trials during FB and TF / good and poor learners based on median split of mean difference between MA of positivity and negativity trials during TF-sessions 14-28	MA negativity trials: FB: session 3-6 < session 25-28 TF: session 3-6 < session 25-28	Difference between MA of positive and negative trials during TF (sessions 14-28) correlates with clinical improvements (hyperactivity/ impulsivity) in good learners
Leins et al., 2007 Group 1 N = 16, Group 2 N = 16	Group 1: Theta↓(↑), Beta↑(↓) (C3f, C4f); Group 2: SCP↑↓ (Cz); 30 sessions, 31-33 follow-up sessions (after 6 months)	Group 1: MP Theta/Beta Group 2: MA of negativity trials. Both: difference between up- und down-regulation	EEG learning both groups: 2nd session < last session 2nd session < follow up	
Bakhshayesh et al., 2011 N = 18	Theta↓/ Beta↑ (Cz); session BL; 30 sessions	MP Theta/Beta across sessions (section 1, 2, 3)	Theta/Beta ↓ in 2 out of 3 training conditions; BL ↓	
DeBeus & Kaiser, 2011 N = 42	Beta↑/(Theta + Alpha)↓ (Fz); 20 sessions	[Beta/(Theta + Alpha)] ↑ (= Engagement Index) of sessions 1-3 compared to 18-20 / Increase of	74 % learners	Teacher rated improvements correlate with change in Engagement Index in

Table 1 (continued)

Liechti et al., 2012 Maurizio et al., 2014 N = 13	Theta↓(↑) / Beta↑(↓); SCP↑↓; tomographic NF of anterior cingulate cortex activity; Pre-session QEEG ; 36 sessions	MP of Beta/Theta or MA across sessions	Only partial learning for a simple SCP variant, otherwise no cross-sessional learning; decrease of pre-session QEEG within-NF-group variability across sessions (normalization)	No association between EEG learning and behavioural outcome, except between SCP delayed feedback regulation and hyperactivity/impulsivity
Hillard et al., 2013 N = 18	Undisclosed protocol (wide band spectrum regulation) (Fpz); 12 sessions	MP frequency analysis at FPz within (minute 1 to 25) and across sessions (session 1 to 12)	Across sessions: Alpha↑ and Beta↓, all other frequency bands↓; Within session: Theta/Beta ↓, Theta/Alpha ↓	
Russell-Chapin et al., 2013 N = 12	SMR↑ (Cz) ; 40 sessions	MP of SMR	SMR↑ (session 1 < session 40)	
Bink et al., 2014 N=45 (adolescents)	Theta↓/SMR↑ (Cz); Session mean 37 (± 5)	MP of Theta/SMR (Alpha, high Beta) of sessions 1 - 5 compared to 31 - 35; Within session first 15 min. compared to last 15 min.	Across session: no change of overall MP; Within-session: Theta↓ larger at sessions 31 - 35 than 1 - 5.	
Escolano et al., 2014 N = 20	Individual upper Alpha↑ (AFz, F3, Fz, F4, FCz and Cz); Pre- and post-session active and passive BL; 18 sessions	MP of individual upper Alpha across sessions and within sessions	Pre-session task-related MP ↑ (= active BL) across sessions; Pre-post MP ↓ within sessions; absolute and relative Alpha MP ↓ within sessions	No association between learning/ training response and behavioral improvements
Gevensleben et al., 2014 N = 10	SCP ↑↓ (Cz); 13 double sessions	MA during positivity or negativity trials / MA↑ across sessions 1, 5, 9 and 13	Cross sessional increase of negative MA during negativity trials	Association between negativity MA of session 5 and 9 and inattention symptoms↓
Takahashi et al., 2014 N=10	SCP ↑↓ (Cz); 16 (20) sessions	Peak amplitude during positivity or negativity trials across sessions	Positive shift amplitude ↑ in session 9, 13; negative shift amplitude ↑ in session 11, 12	
Vollebregt et al., 2014 EEG learning analysed: N = 10	Individualized protocols; most often SMR↑/ Theta↓; 30 sessions	MP per trained frequency-band across sessions	No systematic improvement on target frequencies	

SCP = slow cortical potentials, MA = mean amplitude, MP = mean power, ↓ = decrease, ↑ = increase, TF = transfer condition, FB = feedback condition, BL = baseline, SD = standard deviation

In these NF-SCP studies, learning progress was mostly confined to negativity trials (i.e. to activation), while no or only moderate learning seemed to occur in positivity trials (i.e. deactivation) (Drechsler et al., 2007; Gevensleben et al., 2014; Leins et al., 2007; Strehl et al., 2006, for NF-SCP with healthy adults see Studer et al., 2014). In the initial training sessions, subjects seemed to spontaneously produce positive amplitudes (Drechsler et al., 2007; Strehl et al., 2006), but failed to do so in the subsequent sessions, possibly because they took recourse to more intentional strategies. According to Strehl et al. (2006), children report that the positivity trials are more difficult and exhausting. Alternatively, considering the already high performance in positivity trials during the initial training phase, the lack of improvement in positivity trials might be attributed to a possible ceiling effect (Strehl 2006, Leins 2007). Only one recent study (Takahashi et al., 2014) found comparable increase of positive as well as negative shift amplitudes across training, based on peak amplitudes.

Very few ADHD-studies examined learning of EEG regulation in transfer conditions (Drechsler et al., 2007; Hinterberger et al., 2006; Leins et al., 2007b; Liechti et al., 2012), which is hypothesized to be a more ecologically valid learning measure than performance in feedback trials. In transfer trials, participants regulate their brain activity without feedback or while feedback is delayed. The ability to follow the instructions during transfer trials without the aid of immediate feedback should reflect the child's ability to regulate his/her brain activity independently of external triggers. This ability is considered a necessary precondition for applying the acquired skill in situations outside the laboratory. NF-studies in ADHD reporting learning progress for both transfer and feedback trials are rare (Leins et al., 2007; Table 1). There is evidence that ADHD-subjects are less effective in transfer trials than in feedback trials (Drechsler et al., 2007; Leins et al., 2007; Strehl et al., 2006), which also appears to be the case in

healthy adults (B Kotchoubey et al., 1999). In patients suffering from epilepsy (Lutzenberger, Birbaumer, Elbert, & Rockstroh, 1982; Rockstroh, Elbert, Birbaumer, & Lutzenberger, 1990), EEG regulation performances were comparable in both types of trials.

Several ADHD-studies compared NF-learning to learning progress in other biofeedback modalities, such as muscle relaxation (Bakhshayesh et al., 2011; for a recently published study design see Holtmann et al., 2014b) or biofeedback-guided learning of fine motor skills (Maurizio et al., 2014), with the latter showing better learning with motor than with EEG feedback. Liechti et al. (2012) reported that children with ADHD did not display learning of EEG regulation across sessions in a tomographic EEG NF training. However, they did show progressive learning in muscular artefact control, thus demonstrating a significantly improved ability to sit still.

3.4 Measuring learning of EEG self-regulation

As indicated in Table 1 and illustrated in Figure 1, the methods used for determining the learning of self-regulation with NF-training are heterogeneous. By “learning” (or “EEG-learning”) we will refer to an improvement in a targeted electrophysiological parameter measuring self-regulated brain activity across time, while “EEG training response” implies more generally any training-related change of an electrophysiological parameter (see sections 3.4, 7). We will present a brief overview over different methods and learning indices used in the reviewed studies, discuss possible problems and present additional approaches from studies with other groups than ADHD.

3.5 Units of measurement

The most commonly used units of measurement are the mean level of amplitude and the percentage of time beyond a predefined threshold of EEG activity. The amount of decrease or increase of amplitude in the desired direction or the increased amount of time spent in the desired range of frequencies should reflect the participant's improved regulation efficiency across time. Often, regulation success is dichotomized (yes or no) on each trial, and hit rates are computed online and presented as reinforcers (bonus points) after a block of trials. Such hit rates may be used to represent the EEG learning success across time (e.g. hits above threshold per minute, for children with high functioning autism see Pineda, Carrasco, Datko, Pillen, & Schalles, 2014). This requires, however, that criteria for hits / reward are kept stable, which is not the case with adaptive programs or shaping. Moreover, the use of time units above threshold as criterion is not sensitive to smaller improvements in the regulation of amplitudes just below the threshold.

When considering SCP-NF, the observation of only the change in mean amplitude provides no direct evidence about the participant's ability to differentiate between a state of activation (reflected by a negative amplitude) or deactivation (reflected by a positive amplitude). Nevertheless this skill is hypothesized to be the main training goal in SCP-NF. For the evaluation of progress in learning to differentiate between polarities, it has been common to compute the difference between the means of positive or negative amplitudes and then compare these across sessions (Drechsler et al., 2007; Gani, 2009; Leins et al., 2007b; U. Strehl et al., 2006b). However, this method alone fails to account for cases in which regulation has only been achieved in one direction. To illustrate this, it might be the case that the participant mistakenly produces

an amplitude of moderate negative polarity during the positivity trial, while the performance in the negativity trial is correct (i.e. strong negative polarity) (see Blume, 2012). This objection especially accounts for ADHD-patients, as in several studies cross-sessional learning has been reported only for negativity, but not for positivity trials (Drechsler et al., 2007; Gani, 2009; Leins et al., 2007b; Strehl et al., 2006).

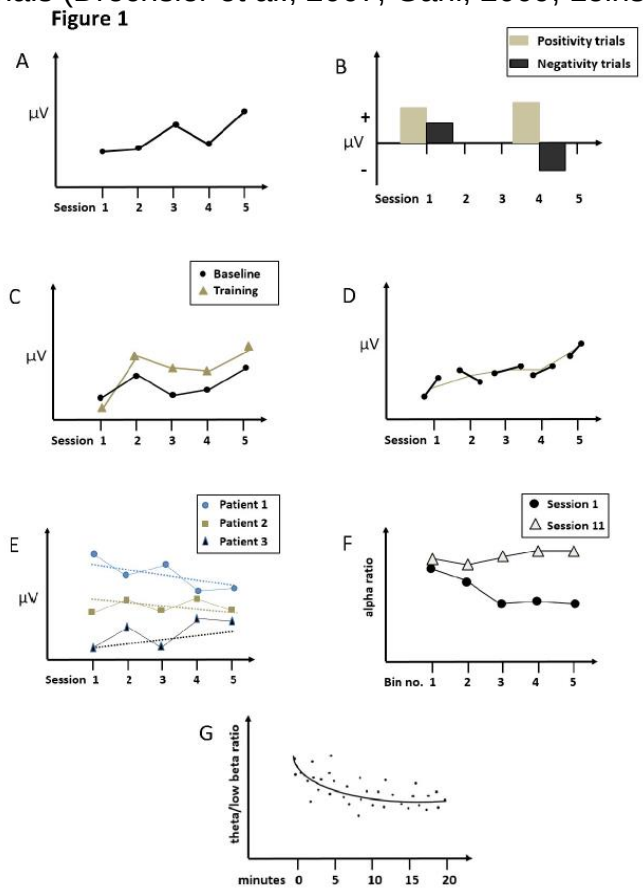

Figure 1. Illustration of across and within session learning curves.A. Across sessions comparison of single sessions (SCP mean amplitude during positivity and negativity trials; adapted from Gevensleben et al. 2014; modified). B. Learning curve across sessions of mean training performance (e.g. Cho et al. 2008; modified). C. Pre-session baseline and mean training performance across sessions (adapted from Dempster & Vernon 2009, modified), D. Pre- and post- session baselines across sessions (adapted from Escolano et al. 2011, modified). E. Individual pre-session baselines across sessions (adapted from Liechti et al., 2012, modified). F. Within session learning curves of training performance during session 1 and session 11, segmented into bins of time (adapted from Cho et al., 2008, modified). G. Within session learning curve collapsed across sessions, indicating mean theta/beta ratio per minute (adapted from Hillard et al. 2013, modified).

3.6 Cross-session learning

In the ADHD studies reviewed, the calculation of cross-session learning was based on different samplings of time periods: several studies used two time periods (session 1 and session 40; Russell-Chaplin et al., 2013) or a small number of selected sessions, usually consisting of one from the beginning, one or two in the middle and one from the end of the training course (1st, 5th, 9th, 13th session: Gevensleben et al., 2014; 1st, 10th. 20th, 30th session: Vollebregt, et al., 2014; see Figure 1 A). However, sampling only a small number of single sessions for the calculation of learning is often problematic as the performance of a single session may be biased due to external variables unrelated to the training (i.e. motivation in the final sessions might be lower, day-to-day events, sleep patterns, etc.). In addition, several studies reported large variability in intra-individual learning performance (Drechsler et al., 2007; Strehl et al., 2006; Leins et al., 2007; for healthy participants e.g. Gruzelier et al., 2014a; Wan et al., 2014). To reduce this large variability throughout the course of the training, some researchers clustered groups of sessions into blocks for analysis, e.g. two sessions into one block (sessions 2/3, sessions 29/30 and follow-up: Strehl et al., 2006; Gani et al., 2008, Leins et al., 2007) or all sessions into three blocks of 10 sessions (Bakhshayesh et al., 2011). Alternately, only the second half of the sessions was incorporated into the (sub-)analysis, as this later phase was thought to be more indicative of learning progress than the first half (Drechsler et al., 2007; epileptic patients: Kotchoubey et al., 1999).

In other studies, training performance has been considered across all sessions, which allows for a more fine-grained analysis of the course of learning also including non-linear changes (Figure 1 B) (ADHD patients: Hillard et al., 2013; Lubar et al., 1995; for

NF learning curves in studies with other clinical groups see e.g. (Enriquez-Geppert et al., 2014; Kouijzer, van Schie, Gerrits, Buitelaar, & de Moor, 2013; Wan, Nan, Vai, & Rosa, 2014). Strehl et al. (2005) argue that a steady learning curve across sessions is not necessary to qualify as a learner, as some subjects might find an optimal strategy only at the end of training.

Large intra-individual variability in cross-sessional EEG regulation performance has also been reported in studies with healthy adults and has been attributed to fluctuating arousal levels. Gruzelier et al. (2014a) refer to healthy participants' self-reported irregularities in night sleep. Indeed, there is evidence that ADHD patients in particular suffer from sleep irregularities (Spruyt & Gozal, 2011). However, the variability of performance due to fluctuations in motivation and arousal is a major feature of ADHD. In order to account for the intra-individual variability of learning performance, Strehl et al. (2006) normalized the data by dividing the individual mean NF-parameters by the individual standard error. This procedure reduces the likelihood of a bias towards subjects with high amplitudes in group analyses of learning. To illustrate this bias, one can imagine a subject with a slow gradual increase in amplitude and thus a small standard deviation. Without normalization, this subject is less likely to reach a predefined criterion of good learning than another subject with a fluctuating pattern.

3.7 Within-session learning

Both within- and cross-session EEG-learning (decrease in theta/low beta and theta/alpha ratios) was reported in ADHD-patients by Hillard et al. (2013), using a wide band EEG regulation training at a prefrontal site. Within-session analyses for theta/low beta ration and theta/alpha ratio resulted in significant decrease in the shape of a logarithmic curve over the 25 minutes of training (for illustration see Figure 1 G). In

addition, significant progressive changes in the expected direction across sessions were found for all analyzed frequencies. Bink et al. (2014) found larger within-session decrease of theta activity during the last sessions of a NF-FR theta/SMR training compared to the first ones, but no significant change of mean power across sessions. Escolano et al. (2014) analyzed within-session learning in an individualized upper alpha training for children with ADHD. Before and after each session QEEGs were recorded with eyes closed (resting EEG, passive baseline) and with eyes open while performing a visual counting task (active baseline). An unexpected pre-post session decrease was found for counting task related EEG activity (alpha “rebound” effect), in contrast to findings by the authors with healthy adults (Escolano et al., 2011).

Different approaches exist to measure within-session learning, e.g. by relating the mean NF-parameters of each period within a session to the first (Wan et al., 2014) or preceding period (Egner & Gruzelier, 2001), collapsed across sessions. Alternatively, a period or a complete session may be divided into very short segments and collapses across sessions (Dempster & Vernon, 2009) or the change of within-session mean parameters may be analyzed across sessions (Cho et al., 2008) (Figure 1 F). Although it might initially seem counterintuitive to examine within-session learning regarding long-term outcome and specificity, there is evidence from NF-studies with healthy individuals that within-session learning collapsed across sessions may be correlated with outcome (Ros et al., 2009). Gruzelier (2014) argues that the consideration of within-session learning would result in a more robust measure of learning than cross-session learning alone, because the overall error variance might be smoothed by a smaller sampling rate of the data within one session averaged over multiple sessions. Several studies with healthy individuals which included both within- and cross-session learning either failed to show cross-sessional NF-learning at all (Hardman et al., 1997;

Cho et al., 2008) or only found a trend (Gruzelier et al., 2014b). By contrast, within-session learning was often evident, i.e. participants improved throughout the session. These findings suggest that it might be interesting to include within-session analyses - or cross-session changes of within-session learning, respectively - more systematically in future NF studies with ADHD.

3.8 Baseline increments

There is increasing evidence from NF studies with healthy adults, that NF may have a strong impact on baseline QEEG, sometimes stronger than on the targeted electrophysiological parameter fed back during the training (Hanslmayr, Sauseng, Doppelmayr, Schabus, & Klimesch, 2005; Ros et al., 2009). As a consequence, EEG-learning should be reflected by a change in pre-session EEG baselines throughout the training course (Gruzelier, 2014). However, only very few NF-studies with ADHD children examined pre-session or pre-post-session changes in EEG spectra. Bakhshayesh et al. (2011) compared session baselines of the first, second and third section of the training and found larger effects for baseline than for feedback parameters. Maurizio et al. (2014, see also Liechti et al., 2013) reported that after combined NF-SCP and NF-FR with tomographic EEG, individual pre-session baseline values gradually converged towards the group mean across sessions, which was interpreted as normalization (Figure 1, E). In an individualized upper alpha-NF for children with ADHD, Escolano et al. (2014) recorded pre- and post-session QEEG and found a significant increase in power across sessions in the targeted parameter in an active pre-session QEEG condition, i.e. when children performed a counting task, while no significant increase in alpha power was obtained either during training or pre-session eyes closed resting EEG.

Several other NF-alpha studies with healthy subjects have shown that by recording a resting-baseline both before and after the training session, the incremental curves constructed from these data provided a more complete picture of the EEG training response over time (Figure 1 D) (Cho et al., 2008; Escolano, Aguilar, & Minguez, 2011; Kouijzer et al., 2013; Zoefel, Huster, & Herrmann, 2011). First, within a training session, the post-session baseline was usually larger than the pre-session baseline. This could be interpreted as a measure of improvement within the session. Second, the overall learning progress achieved during one session was built upon the progress achieved in the previous session. In other words, the baseline measured at the beginning of a session was on the same level as the post-sessional baseline of the previous session. This ratchet-like linear increase in resting baseline seems to indicate that regulation skills are improving throughout the course of the training (Escolano et al., 2011; Figure 1 D). A possible consequence from this finding is that EEG learning across sessions may be masked by progressive increments in resting baseline if these increments are not taken into account in the analysis of change. Compared to the training performance at the first session, target amplitudes may show a cross-sessional increase, even when no increase can be found when considering within-sessional mean amplitudes relative to their respective pre-session baselines (Figure 1 C). Although this remains to be demonstrated for NF with ADHD, NF-alpha-studies with healthy adults lend support to this hypothesis (Dempster & Vernon, 2009). Incorporating a baseline measure might also enhance the comparability of learning performance on group level. For instance, in a NF-study with insomnia patients, Schabus et al. (2014) divided the session mean amplitude of a subject by the corresponding pre-session baseline. As a result, transforming the data into a relative instead of an absolute value may smooth out the high inter-subject variability of baseline measures.

3.9 Classification of good and poor learning

Whereas most of the reviewed ADHD-studies analyze learning improvements of EEG regulation with regard to the full treatment group (Bakhshayesh et al., 2011; Gevensleben et al., 2014; Russell-Chapin et al., 2013; Escolano et al., 2014) some studies report the rate of learners (or responder rate) (Lubar et al., 1995; Kropotov et al., 2005, DeBeus & Kaiser, 2011), or distinguish between good and poor performers (Drechsler et al., 2007) or successful and unsuccessful regulators (Strehl et al., 2006), in order to analyze learning outcome. However, in several NF-ADHD studies which do not include the analysis of EEG learning, the term “responder rate” is used with regard to the clinical outcome, which is usually defined by the reduction in ADHD symptoms (e.g. Gevensleben et al., 2009).

In studies which report the rate of learners, training success may be defined by a fixed criterion, e.g. a percentage cut-off in order to classify participants as learners if they have reached a predefined criterion in a fixed percentage of sessions. These cut-off values for successful learning often appear to be chosen ad hoc (e.g. Kropotov et al., 2005), or may be taken from previous studies (e.g. (Weber, Köberl, Frank, & Doppelmayr, 2011, for NF with healthy adults). In a theta/beta training, Kropotov et al. (2005) defined successful learning by an increase in amplitudes of at least 25 % during feedback periods compared to resting periods in at least 60 % of all sessions. This definition resulted in 82 % participants being classified as “good performers”. The number of training sessions for each patient varied from 15 to 22, depending on several factors such as age, type of ADHD, learning curves, and parent reports. The termination criteria were (1) stabilization of training performance assessed by the

dynamics of the trained parameter during the last three to five sessions, and (2) stabilization of patient's behavior according to parent reports. Lubar et al. (1995) and DeBeus & Kaiser (2011) used a relative change of NF-parameters as a criterion for categorizing performance. In this approach, subjects are classified as good performers when performance in the final training sessions is significantly superior to that in the first ones or when NF parameters increased across all sessions. Lubar et al. (1995) reported a responder rate of 65 % in theta/beta NF-FR training, defined by significant negative correlation of theta by session number. DeBeus and Kaiser (2011) found 74 % of responders in NF-FR training, defined as an increase of half a standard deviation in the Engagement Index (beta/theta + alpha) from session 1-3 to 18-20. (For studies with healthy participants see Dekker et al., 2014; Vernon et al., 2003; Weber et al., 2011; Zoefel et al., 2011).

A different approach is to employ a cut-off value defined by the median split of the learning parameter (Strehl et al., 2006; Drechsler et al., 2007; Doehnert et al., 2008) which allocates the participants into a group of good and a group of poor learners. Naturally, in this case no meaningful responder rate can be given. Moreover, learners and non-learners do not have to be equally distributed, contrary to what the use of median split may lead one to presume. As a consequence, the variability of learning performances may vary considerably in both groups. Evidently, given these methodological differences in the calculation of good learning in the aforementioned studies, it is difficult to draw meaningful conclusions about the average responder rate in ADHD NF. According to a study by Monastra, Monastra, & George (2002), EEG learning essentially appears to be a matter of time. Only children with predefined QEEG abnormalities were included in their study and treatment was continued until the criterion for EEG learning had been obtained in each individual case

(“normalization”, i.e. a degree of cortical slowing within 1.0 SD of age peers). Therefore all participants reached the criterion, which is equivalent to a responder rate of 100 %, but the number of sessions varied considerably among the participants. Further evidence that time may matter with regard to the classification of good and poor learning of EEG self-regulation comes from studies indicating that regulation skills might continue to develop and consolidate long after the end of the training (Blume, 2012; for NF with epilepsy see Strehl et al., 2005).

3.10 Failing to learn

Some studies on NF in ADHD which investigated EEG learning performance failed to find the expected significant changes on group level. In a double blind placebo controlled study using Q-EEG feedback with individualized protocols, Vollebregt et al. (2014) compared mean power of the trained frequency bands of the first, tenth, twentieth and final session. The authors report that seven out of ten children showed changes in power toward the directed target, but no child showed changes in more than one frequency band, and that all children also presented changes away from a training target in some bands. Clinical responders (defined by behavioral improvements) showed EEG changes in both desired and non-desired directions. In a study using tomographic NF, including both NF-SCP and NF-FR, the authors failed to find significant EEG learning on group level (Liechti et al., 2012; Maurizio et al., 2014). Besides methodological aspects, the fact that the regulation of a brain area which is known to be underactivated in ADHD, the anterior cingulate cortex, was fed back, may have presented a special difficulty for the participants. However, in this study patients displayed individual changes towards normalization of pre-session baselines across sessions (Figure 1 E).

Whether or why individual children might fail to learn self-regulation of brain activity has not been the central focus of ADHD-NF research. These questions have been tackled more comprehensively in the Brain Computer Interface (BCI) research, which aims at training individuals to control technical devices via the regulation of brain activity, e.g. to use a communication computer or to navigate a wheelchair controlled by the modulation of brain waves (Guger, Edlinger, Harkam, Niedermayer, & Pfurtscheller, 2003; Vidaurre & Blankertz, 2010). While neurofeedback is based on operant conditioning with a fixed-target EEG signal, BCI most often uses a machine learning approach. This means that the EEG signal is optimized according to the participant's brain activity during the task (Lotte, Larrue, & Mühl, 2013). Nevertheless, a substantial portion of participants, 10 to 30 %, fail to gain control, which has been referred to as BCI "illiteracy" (Dickhaus, Sannelli, Müller, Curio, & Blankertz, 2009) or "inefficiency" (Kübler, Mattia, George, Doron, & Neuper, 2007). Allison and Neuper (2010) presume that a small number of probands may display individual brain structures, which, although not pathological, may not allow the recording of a target EEG parameter by normal surface electrodes (see also Halder et al., 2013). If proper calibration does not help in adapting to individual morphology, the solution is to switch to a different EEG parameter or neuroimaging technology. It is possible, however, that the patient will not be able to use BCI at all. Otherwise, one should try to improve the accuracy of the BCI procedure, e.g. by improving the selection of the existing brain signals through approved algorithms or by incorporating better error correction (Allison & Neuper, 2010). The authors hypothesize that BCI illiteracy might be confined to certain techniques or tasks in a particular individual while the same person may possibly perform better in another paradigm. All of these points are concerned with methodological and technical aspects, while, as the authors state, variables such as mood, motivation, distraction, and test setting may also play a role. In patients with

Amyotrophic Lateral Sclerosis (ALS), motivational factors such as challenge and mastery confidence were positively correlated with BCI performance (Nijboer et al., 2008). However, an exaggerated feeling of self-efficacy may constitute an impediment rather than a help for good NF performance. Witte, Kober, Ninaus, Neuper, & Wood (2013) reported that SMR-learning performance was negatively correlated with the attribution of locus of control. Participants whose confidence in control over a technical device was low performed better than those with a high belief of control. This effect was explained by a possible cognitive overload when controlling a technological device, which in turn might adversely affect the relaxation states which SMR-training aims to achieve. In a study on psychological predictors of SMR learning, the best predictor of SMR performance were objective measures for the accuracy of fine motor skills and the ability to concentrate on the task (Hammer et al., 2012), whereas subjective factors, such as well-being, did not predict performance. This was explained by the fact that only healthy individuals, consisting mostly of students, participated in the study.

To which extent these results from BCI research also apply to NF with ADHD and whether a proportion of children might be unable to learn EEG regulation with one protocol but might gain control with another, is unknown. In future studies, more attention should be paid to the question of whether and why children with ADHD might fail to learn self-regulation of brain activity.

3.11 Learning patterns of self-regulated brain activity

One crucial question is how to interpret patterns of learning curves in terms of learning performance, and whether it is possible to distinguish characteristic learning patterns in ADHD. For the time being, the extent to which the learning of EEG regulation in ADHD may be expected to be progressive and regular remains unclear. Differences in the training administration of ADHD-NF studies (session frequency, time intervals between sessions, number or duration of trials per session, training breaks etc.) and the small number of patients in many studies make it difficult to draw conclusions. For theta/beta-NF, Lubar et al. (1995) (40 sessions) as well as Bakhshayesh et al. (2011) (30 sessions) observed an increment in performance during the first training phase, followed by a stagnation phase in the middle of the training and a subsequent increase in performance in the final third of training sessions. In an SCP-training (Blume, 2012; 25 sessions; 4 weeks-break between session 12 and 13), children with ADHD displayed a stagnation in the second compared to the first training phase, while performance was enhanced again at the 6-months follow-up. Interestingly, some of the children who had been classified as non-learners after the second training phase, showed good EEG performance at follow-up (see Strehl et al., 2014). These learning patterns - stagnation and subsequent increased performance after a break or in the final part of the training - have been discussed in terms of the individual speed of learning and a related overtraining-effect which might occur earlier for fast learners than for slow learners (Blume, 2012). In several studies with healthy participants, NF-FR learning has been reported to reach a plateau after 4-6 sessions with a subsequent stagnation (total session number 8-10) (Enriquez-Geppert et al., 2014; Gruzelier, Inoue, Smart, Steed, & Steffert, 2010; Keizer, Verment, & Hommel, 2010). These plateaus have been hypothesized to reflect training fatigue or over-learning. Patients' learning curve patterns might differ from those of healthy subjects. For instance, Kübler

and Neumann (2004) found that healthy subjects reached a learning plateau after 3 sessions, whereas in patients with ALS, no learning plateau was reached after 12 sessions. In an NF-study with primary insomnia patients, participants displayed fluctuating learning, which, intercepted by sessions of stagnation, increased across sessions (Schabus et al., 2014). In anxiety patients, Hardt and Kamiya (1978) postulated a fifth-order learning curve, starting with an initial increase, and followed by a dip, a second increase, and a final exponential increase for alpha-NF learning.

In healthy individuals, learning curve patterns have been shown to distinguish non-learners from good learners, showing not only a plateau, but also a decrease of performance: Poor SMR performance was associated with a highly significant 10 % decrease in NF-parameters during the second training phase when compared to the first (Ros et al., 2009). A further finding of this study was that smaller intervals between sessions seemed to lead to better EEG learning than longer intervals, indicating that an intense training rhythm may be advantageous.

It should be kept in mind that learning patterns in ADHD besides being extremely individual in nature, may also substantially depend on factors of the setting, such as the relation to the therapist, motivation, external support (Strehl, 2014; Drechsler et al., 2007; Monastra et al., 2002). For the time being, there is a lack of studies that describe characteristic learning patterns and possible subgroups of learners in ADHD which would allow to select the training protocol or to systematically adapt the program according to the learning type of the child.

3.12 Association between self-regulated brain activity and clinical outcome gains

The few studies that examined the association between NF-learning and the clinical outcome in ADHD (see Table 1) used heterogeneous methods. Participants may be categorized in poor and good learners for subsequent data analysis or classified according to good and poor clinical outcome, while in other studies no such distinctions are drawn.

For instance, Strehl et al. (2006) defined criteria for good SCP-learning (negativity learning, calculated by median split) as well as for good clinical outcome in ADHD (at least a 2-point reduction in either hyperactivity or inattention according to DSM-IV) and reported a statistically significant association between the two measures at the end of the training. At the 6-months follow-up, the association between clinical outcome and NF-learning still almost reached significance, indicating a long lasting effect of the training. Drechsler et al. (2007) reported a positive correlation between the pre-post decrease in parent-rated ADHD symptoms and the ability to differentiate between SCP positivity and negativity trials. This association was confined to the group of good performers, defined by median split, whereas in poor learners, ADHD symptomatic improvements were uncorrelated with SCP performance. In NF-FR training, DeBeus and Kaiser (2011) reported a significant correlation between improved EEG regulation and teacher ratings of ADHD symptoms, which was also confined to the group of good performers. Recently, Gevensleben et al. (2014) conducted an SCP-NF study with ADHD children, and found a correlation between the pre-post change in parent-rated inattention symptoms and the increase in negativity from the first to the fifth session and from the first to the ninth session. This study was based on a small sample (n = 10) and the authors did not distinguish between good and poor performers.

Several studies have analyzed the association between EEG learning and neuropsychological outcome. Kropotov et al. (2005) reported that learning to enhance beta and SMR in ADHD correlated with a significant decrease in response time and variability of response time in a Go/No-Go task only for good performers. Lubar et al. (1995) reported stronger improvements on a computerized attention test for learners than for non-learners after NF-FR training.

The relationship between positive clinical outcome and successful NF learning has been confirmed in a number of NF studies with other clinical groups, such as patients with epilepsy (Daum et al., 1993; B Kotchoubey, Blankenhorn, Fröscher, Strehl, & Birbaumer, 1997; Ute Strehl, Kotchoubey, Trevorrow, & Birbaumer, 2005) or sleep disorder (Schabus et al., 2014). In healthy subjects, NF-learning correlated positively with improvement in short-term memory (Nan et al., 2012), mental rotation (Hanslmayr et al., 2005), microsurgical skills (Ros et al., 2009) and enhancement in cognitive creativity (Gruzelier, 2013).

However, it should be kept in mind, that the relationship between successful regulation of an individual's brain activity and positive clinical outcome is not reciprocal: Improvements in parent-rated ADHD symptoms are not confined to learners (Drechsler, et al., 2007), indicating that non-specific treatment effects also contribute to the clinical outcome.

3.13 Electrophysiological pre-post changes, protocol specific effects and prediction

In NF research with ADHD patients, to date no study has directly related pre-post electrophysiological changes to increments in NF performance across sessions.

However, several studies have reported pre-post effects on electrophysiological levels, although most of them did not analyze EEG learning across sessions. Often, these studies focus in a hypothesis-driven manner on electrophysiological measures related to the feedback protocol used, examining pre-post Q-EEG changes after NF-FR with special emphasis on the trained frequency (e.g.(Pop-Jordanova, Markovska-Simoska, & Zorcec, 2005; Thompson & Thompson, 1998) and pre-post contingent negative variation CNV or other ERPs after NF-SCP (Heinrich, Gevensleben, Freisleder, Moll, & Rothenberger, 2004a; Mayer, Wyckoff, Schulz, & Strehl, 2012). There is evidence that training protocols may result in specific effects which, at least indirectly, supports the importance of successful and differential learning of EEG regulation with regard to pre-post EEG changes. Wangler et al. (2011) and Gevensleben et al. (2009) compared NF-SCP and FR-NF training in a crossover design and examined electrophysiological effects of both protocols. They reported pre-post increase in the CNV after NF-SCP but not after NF-FR. According to pre-post QEEG analyses, both protocols resulted in a decrease in theta bands activity. Despite this evidence of protocol-specific effects on EEG, it might be advisable to explore the full frequency spectrum or to include additional measures in the pre-post EEG analyses. Several studies, mostly with healthy participants, demonstrate that electrophysiological pre-post effects are not necessarily confined to the targeted training parameter (for a detailed review see Gruzelier, 2014). An example with ADHD patients is provided by Doehnert et al. (2008) who conducted SCP training and reported a pre/post QEEG theta decrease at Oz, while they did not find the expected effects on the CNV. Another evidence for extended effects comes from a study by Escolano et al. (2014) who in an alpha-NF analyzed the course of pre- and post-session QEEG in resting and in task-related states, though with a focus on the target frequencies. Cross sessional changes in the expected direction were limited to task-related pre-session QEEG while changes in pre-session

resting EEG were not significant. Liechti et al. (2012) were unable to find any significant association between changes in ADHD symptoms and cross-session NF-learning. However, they reported specific associations between cross-session changes in baseline-frequencies and outcome gains, such as a positive correlation between theta/beta increases in specific regions and frontal beta decreases with reductions in hyperactivity/impulsivity. The extent to which in the case of generalized and extended EEG training response the electrophysiological outcome should still be considered the result of a specific training effect should be the subject of a more refined methodological debate.

Electrophysiological pre-post changes have been related to clinical outcome, which indicates that electrophysiological change is reflected by behavioural improvement (Arns, Drinkenburg, & Leon Kenemans, 2012; Doehnert, Brandeis, Straub, Steinhausen, & Drechsler, 2008; Gevensleben et al., 2009b; Wangler et al., 2011). Still, electrophysiological pre-post measures do not directly reflect EEG regulation performance during feedback trials. Pre-post changes in electrophysiological markers have also been reported after mindfulness training (Moore, Gruber, Derose, & Malinowski, 2012; Schoenberg et al., 2014), which shares several therapeutic characteristics with the NF setting, and thus results based on these measures do not provide the best indication of NF specificity.

Studies that analyze initial EEG learning patterns across or within sessions with regard to overall EEG learning performance, are rare. However, the identification of early predictors of NF learning would be very helpful in terms of providing a better basis for therapeutic decision-making or adapting the training protocol accordingly. In an unpublished doctoral thesis by Goth (2006) on NF training in children with ADHD, the mean amplitudes of negativity trials in session 1 and 2 were the best predictors of

subsequent improvements in SCP-NF-regulation performance, whereas a large number of inattention symptoms predicted poor EEG learning. In NF-FR training, a similar trend was found for successful regulation in early sessions. The best predictor of EEG learning success in NF-FR, however, was a high IQ.

In patients with ALS, good performance at an early training stage of SCP regulation was correlated with subsequent good learning (Neumann & Birbaumer, 2003). In a study with healthy adults, it could be shown that certain morphological parameters may have a beneficial effect on training success: Frontal-midline theta NF-learning was predicted by the volume of the mid-cingulate cortex and the white matter concentration of underlying brain structures (Enriquez-Geppert et al., 2013).

3.14 Is it possible to promote EEG self-regulation performance?

It has been suggested that children with ADHD might require explicit rather than implicit learning (Lansbergen, van Dongen-Boomsma, Buitelaar, & Slaats-Willemse, 2011). According to several authors in the field, children with ADHD need to actively practice mental strategies to self-regulate brain activity and have to be instructed on how to translate the newly learned skill into everyday life (Gevensleben et al., 2009; Heinrich & Gevensleben, 2013; see Strehl, 2014). They suggest that during the first lessons of training, the trainer should encourage the child to find an appropriate strategy ("I imagine I'm waiting for the starting signal in a race"). This initial strategy should be gradually reduced and finally abandoned in the course of the training, when regulation becomes automatized (Heinrich et al., 2004a). To the best of our knowledge the impact of instruction and explicit strategy training on EEG training performance has not been systematically investigated in ADHD. Gevensleben et al. (2014) hypothesize that the use of different transfer instructions for children with Tic disorder than for children with

ADHD may have resulted in specific clinical outcome gains in inhibitory control. However, these setting differences did not apply to the self – regulation during feedback trials, but to the transfer outside the laboratory. Whether self-regulation of brain activity may be helped or exacerbated by the use of conscious top-down strategies is unclear and probably also depends on specific protocols. As SCP training aims at quick changes in polarity, it may be expected that top-down regulation plays a more prominent role here than in NF-FR (see Loo & Makeig, 2012). Arguments both for and against the promotion of conscious strategy use and the importance of self-awareness for NF performance come from research with healthy subjects and other clinical groups. Neurofeedback has been hypothesized by several researchers to involve an increased awareness of the physiological states underlying the feedback (Cannon et al., 2007; Plotkin & Rice, 1981). Recent evidence for this hypothesis is provided by a study on EEG discrimination training with healthy adults (Frederick, 2012). After a baseline recording (150 secs.), subjects had to respond to a prompt asking whether in that moment they were in a low (< 30th percentile of the baseline) or high alpha state (> 70th percentile). They immediately received feedback about their guess. 75 % of participants showed a significant learning curve and were successful in discriminating their brain activity states. There might be a reciprocal relationship between discrimination of brain states and the training of brain state regulation, as Cinciripini (1984) showed for SMR and (Boris Kotchoubey, Kübler, Strehl, Flor, & Birbaumer, 2002 (20 demonstrated for SCP-training. Moreover, successful regulation skills might also have a positive impact on the discrimination ability of brain regulation states. Gruzelier (2013) reports that the subjects' first positive self-judgment about their ability to regulate SMR ratios occurred close to the time, when their learning curve reached a plateau.

A further question concerns whether and how mental strategies might affect NF-learning. Nan et al. (2012) reported that their (healthy adult) alpha-NF subjects favored positive mental strategies (e.g. friends, love, family) which they estimated the most successful. However, these subjective judgments were not related to the actual NF-performance. The effects of strategy use might also depend on the frequency band: in NF-SMR training with healthy adults, participants who used no mental strategy at the end of the training performed better than those who did, thus indicating a possibly counterproductive effect of strategy use on SMR learning. In contrast, strategy use had no influence on gamma learning (Kober, Witte, Ninaus, Neuper, & Wood, 2013). Neumann and Birbaumer (2003) argue that providing patients with initial strategies may promote self-regulation at the beginning of training but would prevent subjects from trying out other potentially more effective strategies with training progress. This argument is in line with Witte et al. (2013) who emphasize the importance of the initial trial-and-error learning, which due to “immediate closed-loop feedback” could ameliorate the subjects’ regulation skills. This unconscious adapting to the desired state might thereby become automated.

To conclude, the literature provides arguments both against and in favor of a more systematic approach to foster EEG learning and self-awareness of EEG activity states in children with ADHD. It might be worthwhile to devote more attention to the question of whether and how the learning of EEG self-regulation can be systematically promoted in children with ADHD.

3.15 Conclusion

Discussions about NF specificity need to include analyses of EEG regulation performance and its impact on clinical outcome. Besides its effects on ADHD primary

symptoms, associations with factors usually regarded as “generic effects”, such as improved self-perception or self-efficacy should also be considered. To provide optimal conditions for learning, it is necessary to improve our knowledge regarding characteristic cross-session learning trajectories and within-session performance in ADHD and to adapt training schedules accordingly. This also includes possible therapeutic strategies which might promote EEG self-regulation in children with ADHD. In the future, NF devices used for NF research with ADHD should adhere to more rigorous scientific standards, allowing for qualitatively acceptable EEG recording during treatment sessions, including artefact control. From a scientific point of view, the current practice, which allows the use of NF devices of uncertain quality or protocols based on undisclosed algorithms for NF research, is unsatisfactory. It is bewildering that, with regard to the evaluation of efficacy and specificity of NF, strictest methodological standards are demanded for the study design, while no scientific standards need to be applied to the treatment. Several meta-studies (Martijn Arns, Heinrich, & Strehl, 2014; Hodgson, Hutchinson, & Denson, 2014) have demonstrated the efficacy of NF with regard to the improvement of ADHD symptoms. Whether NF is efficacious AND specific still needs further investigation, which should go beyond analyzing pre-post changes and include analyses of the treatment process and the learning of EEG self-regulation.

4 Paper 2 - Mixed-effects modeling of Neurofeedback-training performance in children and adolescents with Attention-Deficit/Hyperactivity Disorder – predictors for learning

Agnieszka Zuberer[1], Franziska Minder[1], Daniel Brandeis[1234], Renate Drechsler[1]

[1] Department of Child and Adolescent Psychiatry, University of Zurich, Zurich, Switzerland
[2] Neuroscience Center Zurich, University of Zurich and ETH Zurich, Switzerland
[3] Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, [4]Medical Faculty Mannheim/ Heidelberg University, Mannheim, Germany

4.1 Abstract

Introduction

Neurofeedback (NF) has gained increasing popularity as a training method for children and adults with Attention-Deficit/Hyperactivity Disorder (ADHD) and as a possible alternative to stimulant treatment. However, little is known about the learning that takes place during the training. It is unclear whether and how children learn to regulate their brain activity and whether NF learning is affected by subject- (e.g. clinical symptoms, IQ, age) or treatment- (e.g. setting, inter-session interval, combination with stimulants) related aspects.

Methods

In total, 48 subjects (age 8.5-16.5 years) with ADHD underwent 15 double training sessions of slow cortical potential NF either in a clinical or school setting. Of these, 17 were on constant methylphenidate medication. A mixed effects modeling approach

analyzed within- and cross-session NF learning for two different conditions (Feedback and Transfer).

Results

Being on constant stimulant medication and increased age were associated with stronger cross-session NF learning when performance feedback was provided. In contrast, when no performance feedback was provided, only being on constant stimulant medication appeared to be associated with NF learning across sessions. Except for IQ, other subject-related (e.g. clinical ADHD symptoms at screening, sex) or treatment-related factors (e.g. inter-session time interval, setting) did not improve the model fit.

Conclusions

This first study analyzing predictors of NF learning in ADHD with a mixed-effects approach revealed that NF learning in older children improved with stimulant medication when performance feedback was provided. Although no other clinical or training-related predictors could be identified, the approach may benefit future analyses of NF learning.

4.2 Introduction

Neurofeedback (NF) is a training method by which real-time feedback of brain activity, typically an EEG parameter, is delivered to the subject to promote voluntary control of brain activity and consequently of mental states. The subject has electrodes attached to the head, and the measured EEG parameter is converted to a sound or visual stimulus, which is then fed back to the subject. The main NF protocols for patients with Attention-Deficit/Hyperactivity Disorder (ADHD) are the training of frequency bands and the training of slow cortical potentials (SCPs). Frequency-band NF targets tonic aspects of activation by promoting learning to reduce or to enhance activity of defined frequency bands. SCP training targets the phasic regulation of cortical excitability by learning to generate negative and positive shifts of cortical activity. SCPs originate in the apical dendritic layers of the neocortex and reflect synchronized depolarization of large groups of neuronal assemblies. According to Birbaumer's threshold regulation model of cortical excitation (Birbaumer et al., 1990), negative and positive SCPs are associated with an activated or deactivated state, respectively.

Although frequency band training is the most common form of NF for ADHD, recent research no longer supports the presumption that increases in theta power, reductions in beta power, or the theta/beta ratio are reliable ADHD markers and, in consequence, compelling targets for NF (Arns, Conners, & Kraemer, 2013; Arns et al., 2012; Doehnert et al., 2008) (M. Arns et al., 2013; Cerquera, Arns, Buitrago, Gutiérrez, & Freund, 2012). A rationale for using SCP NF is the relatively robust finding of an ADHD-related reduction of the contingent negative variation (CNV), a SCP which reflects preparation and activation and has been shown to normalize partially after SPC-NF training (e.g.Wangler et al., 2011); Heinrich, Gevensleben, Freisleder, Moll, &

Rothenberger, 2004b). In addition, regulation rather than normalization may be the target of the training (Holtmann et al., 2014).

In recent meta-analyses of NF efficacy for ADHD (Cortese et al., 2016; Sonuga-Barke, Brandeis, Holtmann, & Cortese, 2014), significant treatment effects were found for parents', but not for teachers' ratings ("probably blinded"). These reviews did not consider whether subjects showed successful NF learning; however, this is an important aspect of training specificity. If children show good clinical improvements without successful NF learning, changes have to result from other, nonspecific aspects of the training. We will refer to the ability to modulate the NF parameter as "NF learning" without presumptions about its efficacy (in line with e.g. (De Zambotti, Bianchin, Magazzini, Gnesato, & Angrilli, 2012; Nan, Wan, Vai, & Da Rosa, 2015). The few studies that have examined NF learning across the course of the training differed considerably in their methodological approaches and definition of learner rates (see Zuberer, Brandeis, & Drechsler, 2015, for a review).

In addition, it has been argued that the analysis of within-session learning would result in a more robust measure than analyzing cross-session learning alone (Gruzelier, 2013). From a clinical perspective, such within-session analysis also allows progressive fatigue effects towards the end of a session to be controlled for.

Transfer trials, wherein the subject has to modulate the NF parameter without the aid of an immediate feedback stimulus, are hypothesized to be a more ecologically valid learning measure than the training condition with direct feedback (Strehl et al., 2006b), but few studies report results on that type of task (Drechsler et al., 2007; Leins et al., 2007a; Liechti et al., 2012; Strehl et al., 2006a).

Improving the clinical efficacy of NF and permitting a better informed treatment allocation both require that the individual predictors of NF learning are identified.

However, these predictors have rarely been investigated, and the few analyses that did so, did not lead to conclusive results (e.g. Drechsler et al., 2007). This lack might be due in part to small sample sizes and the ensuing limitations of traditional methodological approaches.

Neurofeedback for ADHD has mainly been perceived as an alternative for stimulant medication, but the combined effects of medication on NF learning are unknown. In several ADHD NF studies, stimulant intake has been permitted in constant dose (Bakhshayesh et al., 2011; Drechsler et al., 2007; Liechti et al., 2012) or without explicit restrictions (Gevensleben, Kleemeyer, et al., 2014b; Strehl et al., 2006a; Vollebregt et al., 2014), whereas in other studies, it has been an exclusion criterion (Escolano et al., 2014; Kropotov et al., 2005). Moreover, the few studies that included medication effects in their analyses did not consider their impact on NF learning (Monastra et al., 2002; Strehl et al., 2006b, Bink et al., 2014). There is little evidence on how stimulants might affect NF learning in ADHD (Clarke, Barry, McCarthy, Selikowitz, & Brown, 2002; S. K. Loo, Teale, & Reite, 1999; J. N. Swartwood, Swartwood, Lubar, & Timmermann, 2003; M. O. Swartwood et al., 1998). Further, although a great deal of evidence suggests that EEG activity is associated with age (Clarke, Barry, McCarthy, & Selikowitz, 2001; Matousek & Petersén, 1983; Matthis, Scheffner, & Benninger, 1981), to our knowledge, it has not been employed as a possible covariate for NF learning. It is also unknown whether contextual and administration factors, such as intensity and duration of sessions, training location, and context—for instance at school, in a summer camp, or in a clinical setting—may systematically alter the ability to regulate one's brain activity.

NF learning in ADHD has rarely been analyzed by considering both inter- and intra-subject variability across time, which would allow for correlation between observations within both a unit and groups, as is possible with mixed-effects models (see Baayen, Davidson, & Bates, 2008). One major advantage of this statistical approach is that it does not assume independence among observations. In addition, mixed-effect modeling is to some degree more robust with unbalanced data than multivariate analysis.

In this paper, we tackle three questions: 1. how both subject-specific and treatment-related factors might be related to NF learning within and across sessions; 2. whether results differ in Feedback and Transfer conditions, and 3. whether and how within-session analysis can contribute information additional to cross-session analysis.

4.3 Methods and Materials

4.3.1 Participants

Subjects were recruited in outpatient clinics, by referral of clinicians, in parent self-aid groups, and at schools. Forty-four subjects were included in the study. See Table 1 for group characteristics.

Table 1

Description of participants

N (total)	48
Male/female (N)	28/20
Setting school/clinic (N)	22/26
Mean inter-session interval (days)	
Clinic	4.1 ± 1.9
School	4.8 ± 1.1
Age (years) all	11.2 ± 2.2
on MPH	11.19 ± 2.4
without MPH	11.09 ± 2.0
Methylphenidate (with/without)	17/31
Dosage (mg)	24.5 ± 15.1
Duration intake (years)	2.3 ± 2.5
Estimated IQ	109.5± 14.8
DSM-IV Conners-3 Parents (T-scores)	
Inattention	67.8 ± 5.8
Hyperactivity / Impulsivity	64.9 ± 8.4
DSM-IV Conners-3 Teacher (T- scores)	
Inattention	65.7 ± 6.1
Hyperactivity / Impulsivity	63.0 ± 7.9

Inclusion in the study required written consent by both the child and parents. The study was approved by the local ethics committee. Age ranged from 8.5 to 16.5 years. Inclusion in the study was based on clinically relevant scores in the German version of the Conners-3 Parent and Conners-3 Teacher Rating Scales (Lidzba, Christiansen,

Drechsler, 2013), according to DSM-IV criteria (one of two ADHD DSM- IV indices reaching T-values ≥ 65, the other T ≥ 60 according to both teachers' and parents' ratings for children of the combined subtype; ADHD DSM-IV inattention T ≥ 65 in one and T ≥ 60 in the rating for the inattentive subtype).

Medication with methylphenidate (MPH) was allowed if the dose was kept stable over the full treatment time, including three months before the first assessment. For children taking MPH, teacher and parent ratings had to be based on the behavior on medication. Exclusion criteria were estimated IQ ≤ 80 (short form of the German WISC-IV; Waldmann, 2008), taking atomoxetine or a neuroleptic or other psychoactive drug, severe comorbidities or other psychiatric disorders, neurological disorders, previous experience with NF (more than four lessons), or either participating in or planning to start a treatment which might confound training effects. Sufficient knowledge of the German language was a further precondition so as to fully understand instructions (children) or to complete questionnaires (parents). Parents had to complete the Development and Well-Being Assessment (DAWBA; (Goodman, Ford, Richards, Gatward, & Meltzer, 2000) to screen for comorbid clinical conditions.

4.3.2 Study design

Parents and teachers rated the child's behavior on the Conners-3 scales and the Behaviour Rating Inventory of Executive Function (BRIEF) (Gioia, Isquith, Guy, & Kenworthy, 2000) at three time points: 3 months before training onset, directly before training onset, and directly after training end. This study focusses on the NF treatment phase of a larger project that involved additional assessments and another treatment group. Their specifications are not relevant for the present analyses and are described elsewhere.

About half of the children (N=23) underwent NF training in the outpatient clinic of the Department of Child and Adolescent Psychiatry (clinical setting), and the other children (N=21) were trained at school in a separate room (school setting).

A complete training comprised 15 double sessions (120min) administered over 10 to 12 weeks. In the clinical setting, training started as a 2-week vacation course with double training sessions daily (five double sessions per week; see Fig. 1) followed by weekly double sessions over at least five weeks. A maximal break of 10 days was permitted during the last training phase (e.g. during vacation). In the school setting, two to three sessions per week were administered for the first two weeks, followed by one weekly session over at least seven weeks (see Fig. 1).

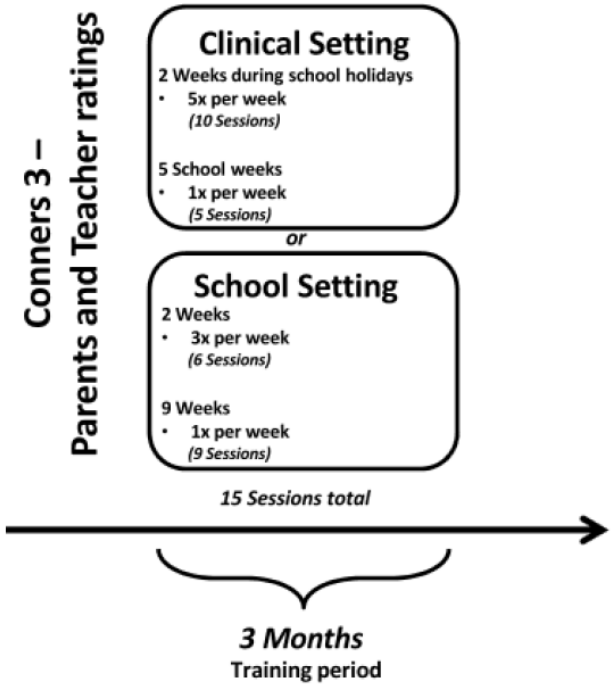

Figure 1. Study design of neurofeedback training.

4.3.3 Description of the NF training

NF was provided using a commercially available mobile training device (Theraprax; NeuroConn GmbH). The subject was seated in a comfortable chair in front of a computer monitor. The NF training was presented as a computer game. Depending on the color and direction of a centrally fixated triangle, the subject was instructed to either activate (produce negative SCP shifts; red upwards-pointing triangle) or deactivate (produce positive SCP shifts; blue downward-pointing triangle). This task was performed either with or without a direct feedback stimulus (Feedback or Transfer condition, respectively). In the Feedback condition, the subject was instructed to steer a stimulus (e.g. fish, airplane) above or below a central horizontal line while it moved from left to right across the screen. The change in activation was fed back by the target stimulus, whose vertical position was proportional to the SCP shift. Good performance

(stimulus was kept at least two seconds above or below a predefined threshold of ±40μv) was rewarded in both conditions by a reward stimulus (sun) at the end of the trial. All conditions and tasks appeared in randomized order (see Fig. 2). The proportion of activation and deactivation trials was always equal (50% each). The percentage of transfer trials increased gradually with session and block number (percentage of transfer trials per double session: double sessions 1-2: 20/ 20/ 20/ 20, double sessions 3-5: 20/ 20/ 20/ 40, double sessions 6-8: 20/ 20/ 40/ 40, double sessions 9-13: 20/ 40/ 40/ 50, and sessions 14-15: 50/ 50/ 50/ 50). One training day consisted of two training sessions of 60 minutes each separated by a break of five minutes. Double sessions consisted of four blocks, each containing 40 trials (see Fig. 2).

4.3.4 Montage and EEG recording

The participants' EEGs were recorded at electrode Cz, referred to two mastoid electrodes shunted over a 10-kOhm resistance (Impedances < 20 kOhm; sampling rate was 512 Hz). The EEG amplifier (Theraprax Neuroconn©) used a low-pass filter of 40 Hz. Processing of the slow cortical potentials (DC - 2 Hz) including eye-movement correction was performed from channel Cz-A2 for each sample point and displayed on the trainer screen. The maximal time delay until the patient saw the feedback of the NF parameter was about 110 ms. Display of the change in mean amplitude was fed back by the horizontal movement of the feedback stimulus, whereas its vertical position corresponded to the difference in amplitude with respect to the recorded baseline of 2 seconds, recorded before trial onset. One SCP trial lasted 8 seconds and consisted of three phases (see Fig. 2): A baseline phase (seconds 0–2), an active phase (8s), and a reinforcement phase (2s). The SCP in the active trial of the last 6 seconds was corrected by the pre-trial baseline (see Fig. 2) and then averaged. Filtering of NF

learning was performed from 0.01- 40Hz with a two-way least-squares FIR filter. Pre-processing was performed with MATLAB and EEGLAB.

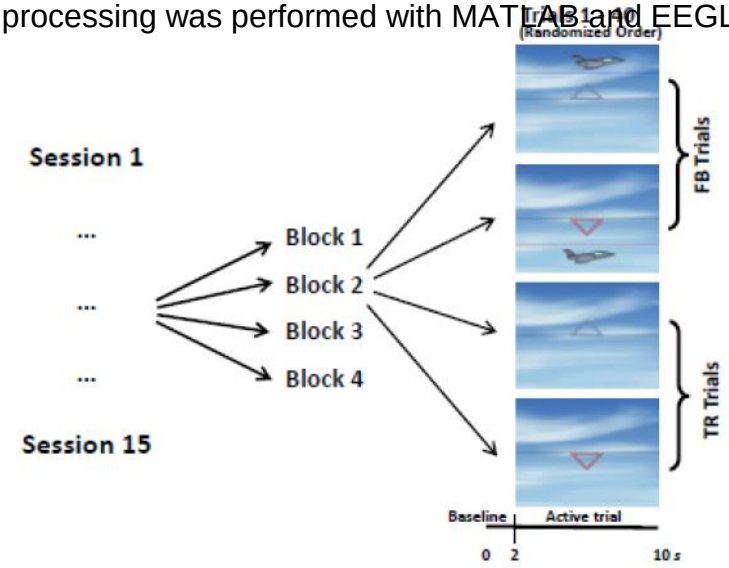

Figure 2. Trial setup of neurofeedback training.

4.3.5 Statistical analysis

Statistical analysis was performed with a linear mixed-effects (LME) regression (Baayen et al., 2008). A random effect was retained if there was a significant difference between the log-likelihood ratio of a model that contained the random effect and a model that did not ($p < 0.05$). Following the principle of marginality, main effects for higher-order interactions were kept in the model (DebRoy & Bates, 2003). To control for high Type I error rate inflation, we also included a random slope coefficient in the model (Baayen, 2008; Barr, Levy, Scheepers, & Tily, 2013). Statistical analysis was performed using the lme4 package in R (D. M. Bates, Maechler, Bolker, & Walker, 2015).

The dependent variable was mean amplitude (µV). For cross-session analysis, the mean amplitude of each baseline-corrected trial was averaged for each session. For within-session analysis, the mean amplitude of each baseline-corrected trial was

averaged across sessions and then further averaged across 10 equally spaced units (from here on called bins). All analysed effects are summarized in Table 2.

Table 2

Effects considered for statistical analysis

Model specifications	***Measure***
Time	
Cross-sessions model	Double session number (15 double sessions).
Within-session model	Bin number; 10 bins per session. The mean amplitude of baseline-corrected trials were averaged across-sessions, and then averaged across the ten equally spaced units (bins).
Task Type	
Feedback (FB)	Online feedback stimulus visible.
Transfer (TR)	No online feedback.
Effects	
Conditions	Deactivation (generation of positive potential shifts of SCPs) vs. Activation (generation of negative potential shifts of SCPs).
Inter-session- interval	Days passed between training sessions.
Age	In years (continuous variable in the model. Only for visualization in plots dichotomized into younger and older age classes).
MPH	Being on constant stimulant medication (methylphenidate), factorized into yes versus no.
Stimulants intake duration	Years of MPH intake.
Dosage of stimulant medication	Methylphenidate (MPH) in mg.
Sex	Factorized into female versus male.
IQ	Estimated IQ (WISC-IV short form).
Setting	Factorized into school setting versus clinical setting.
Severity of ADHD symptoms	T-values of the Conners-3 DSM-IV indices for Hyperactivity and Inattention based on parent and teacher ratings at screening.
Pre-existing ADHD diagnosis	Clinical ADHD diagnosis before entering the study factorized into yes versus no.
Artifact rate	Percentage of rejected trials within a session.

4.4 Results

4.4.1 Cross-session NF learning

The results of the four model analyses with respect to condition type (feedback and transfer) and time (within-session NF learning) are summarized in Table 3.

Table 3

Results for NF learning with respect to condition (Feedback/Transfer) and time (cross-/within- session)

	Cross-session learning						Within-session learning					
	Feedback			Transfer			Feedback			Transfer		
	B	CI	p	B	CI	p	B	CI	p	B	CI	p
Fixed Parts												
Intercept	-3.95	-5.58 – -2.44	**<.001**	-1.06	-2.72 – 0.68	.218	-2.65	-3.86 – -1.40	**<.001**	1.27	-0.38 – 2. 68	.095
Session	-0.00	-0.17 – 0.16	.994	0.03	-0.18 – 0.21	.797						
Bins							-0.29	-0.40 – -0.18	**<.001**	-0.40	-0.61 – -0.20	**<.001**
Task	1.66	-0.04 – 3.31	.061	-2.40	-4.50 – -0.56	**.014**	2.79	2.14 – 3.41	**<.001**	-2.92	-4.34 – -1.27	**<.001**
Age	-0.49	-1.28 – 0.27	.214	-0.59	-0.94 – -0.24	**.002**	-0.62	-1.20 – -0.16	**.014**	-0.40	-0.78 – -0.00	**.040**
MPH	1.18	-1.44 – 3.66	.384	4.11	1.34 – 7.13	**.006**	1.17	-0.57 – 2.87	.176	1.71	-0.03 – 3.40	.051
IQ	-0.08	-0.14 – -0.02	**.006**				-0.07	-0.13 – -0.01	**.018**			
Session : Task	0.15	-0.02 – 0.34	.123	0.20	-0.02 – 0.4	.072						
Bins : Task										0.33	0.07 – 0.56	**.011**
Session : Age	-0.01	-0.09 – 0.08	.858									
Task : Age	-0.42	-1.33 – 0.43	.342				-0.40	-0.71 – -0.08	**.016**			
Session : MPH	-0.05	-0.32 – 0.21	.745	-0.27	-0.63 – 0.06	.124						
Task : MPH	-1.56	-4.53 – 1.91	.308	-4.35	-7.54 – -1.13	**.010**	-0.79	-1.98 – 0.39	.177			
Age : MPH	0.38	-0.90 – 1.67	.538				-0.73	-1.45 – 0.09	.066			
Session : Task : Age	-0.01	-0.12 – 0.09	.804									
Session : Task : MPH	0.12	-0.20 – 0.46	.464	0.39	0.05 – 0.76	**.036**						
Session : Age : MPH	-0.15	-0.28 – -0.01	**.029**									
Task : Age : MPH	-1.41	-2.70 – 0.019	**.042**				0.86	0.36 – 1.35	**.001**			
Session : Task : Age : MPH	0.32	0.16 – 0.47	**<.001**									
Random Parts												
σ^2	41.776			49.174			17.870			32.477		
$\tau_{00,\ subject}$	7.214			8.501			9.093			13.320		
$\tau_{11,\ session}$	0.0734			0.1325								
$\tau_{11,\ bins}$							0.05569			0.1231		
ρ_{01}	-0.455			-0.620			-0.585			-0.735		
Observations	1400			1380			959			959		

Note. Visualization of cross- session NF-learning in the Feedback condition. The dependent variable is mean amplitude (µV) of baseline corrected trials. *Feedback/Transfer*: condition where a feedback stimulus is (Feedback) or is not (Transfer) visible. *Session*: session number. 15 double training sessions in total. *Bins*: bin number; 10 bins in total. *Task:* performance in the deactivation (generation of positive potential shifts) versus activation task (generation of negative potential shifts). *MPH*: on constant methylphenidate medication (yes versus no). σ^2: within-subject residual variance. $t_{00,\ subject}$: between-subject variance. t_{11}: random-slope variance. ρ_{01}: random intercept-slope correlation.

As shown in Table 3, the final model for cross-session learning for the Feedback condition included subject as random intercept (τ00= 7.214) and session number as random slope (τ11=0.0734). A four-way interaction between session number, task, age, and stimulant intake resulted in the best model fit (β = 0.32; CI = 0.16 – 0.47; p< .001). A higher IQ was associated with a more negative mean amplitude (β = -0.08; CI = -0.14 – -0.02; p= .006). The inclusion of the remaining effects summarized in Table 2 did not result in a better model fit.

As shown in Figure 3, the desired learning pattern, showing a positive slope in the deactivation task and a negative slope in the activation task, became more prominent with increasing age and stimulant intake. To test for possible over-parametrization effects due to the complex four-way interaction, a separate model was analyzed in which the task effect was omitted and accounted for in the dependent variable: The dependent variable was the SCP differentiation, the difference between mean amplitudes of deactivation and activation. The results are in line with the original model (see S1 and S2).

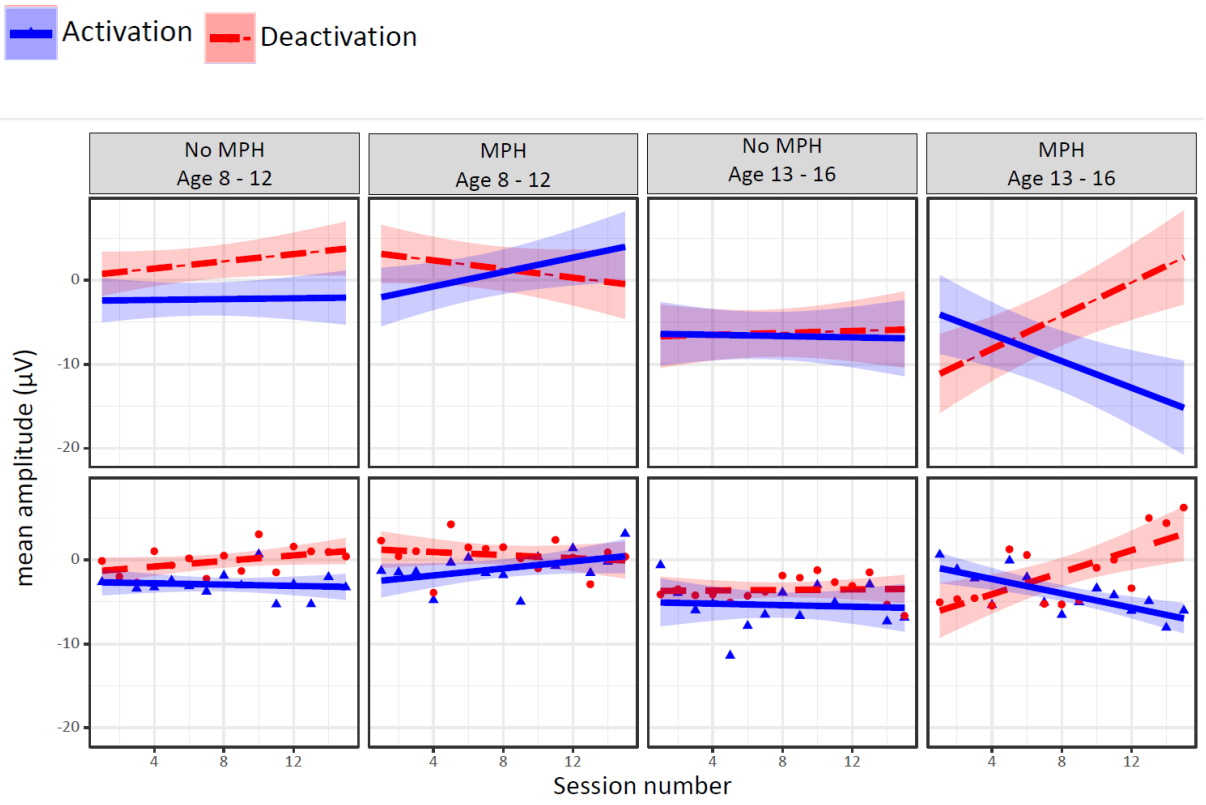

Figure 3. Visualization of cross-session NF-learning in the Feedback condition. The dependent variable is mean amplitude (µV) of baseline corrected trials. For comparison between effects and raw data, see scatter plot under each effects panel, fitted with a fixed linear regression based on the same factors as in the effect plots. Interaction plot for the fixed effects session number, Task (deactivation versus activation), MPH and age. Session number: 15 sessions in total. Deactivation: Generation of positive potential shifts. Activation: Generation of negative potential shifts. MPH: being on methylphenidate medication (yes versus no). For visualization age is subdivided into two age classes (8-12 and 13-16 years), but preserved as a continuous variable in the original model. For the Transfer condition, MPH intake duration was factorized into the levels being on constant stimulation (yes versus no), but preserved as a continuous variable in the original model.

As shown in Figure 4, the final model for cross-session learning for the Transfer condition included subject as random intercept (τ00= 7.860) and session number as random slope (τ11=0.1154; Table 3). A three-way interaction between the fixed effects session number, task, and MPH resulted in the best model fit (β = 0.39; CI = 0.05 – 0.76; p= .036). Higher age was associated with a more negative mean amplitude (β = -0.59; CI = -0.94 – -0.24; p= .002). Thus, NF learning was rather prominent when being on constant methylphenidate medication. The inclusion of effects of the remaining factors summarized in Table 2 did not result in a better model fit.

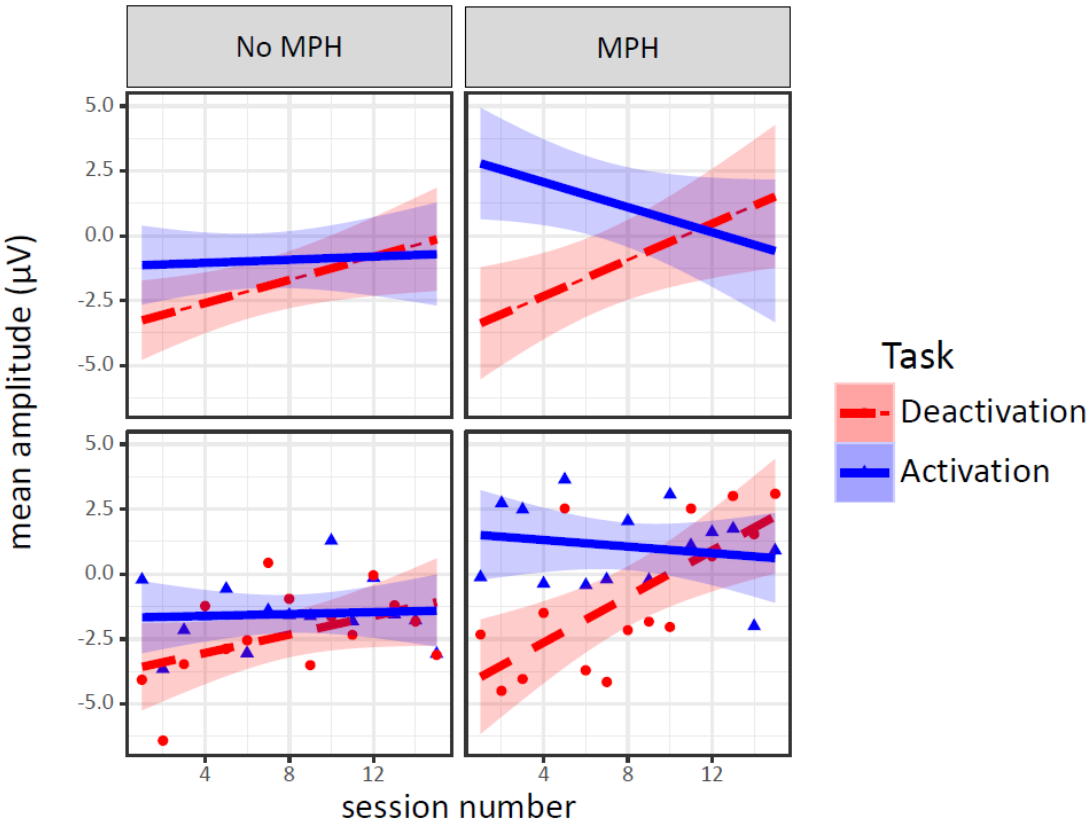

Figure 4. Visualization of cross-session NF-learning in the Transfer Condition. Interaction plot for the fixed effects session number, task and MPH). For comparison between effects and raw data, see scatter plot under each effects panel, fitted with a fixed linear regression based on the same factors as in the effect plots. For visualization, MPH was factorized into two levels (yes versus no), but preserved as a continuous variable in the original model.

4.4.2 Within-session NF learning

4.4.2.1 Feedback condition

The final model for within-session learning for the Feedback condition included subject as random intercept (τ00= 9.093) and bin number as random slope (τ11= 0.05569). As shown in Figure 5, an increasing bin number was associated with a more negative mean amplitude (bins: β = -0.29; CI = -0.40 – -0.18; p< .001). A higher IQ was associated with a more negative mean amplitude (β = -0.07; CI = -0.13 – -0.01; p= .018).

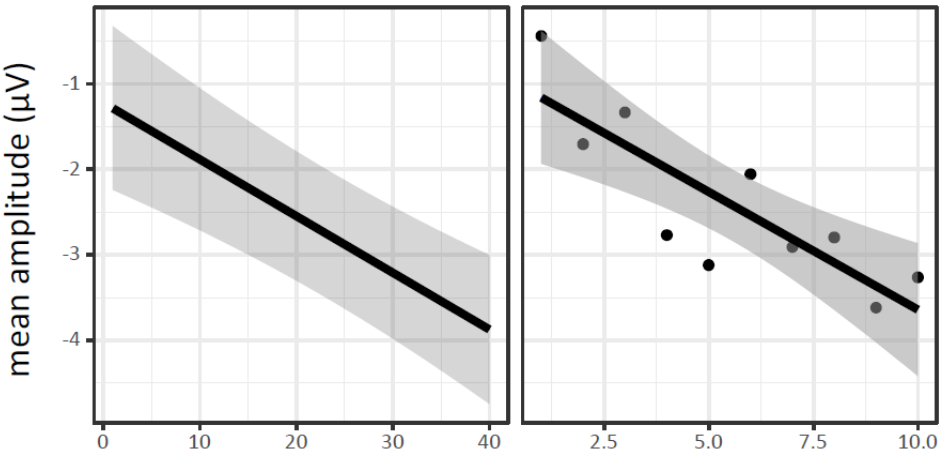

Figure 5. Visualization of within session-NF-learning in the Feedback Condition. The dependent variable is mean amplitude (µV) of baseline corrected trials. Interaction plot for the fixed effect bin number. For comparison between effect and raw data, see scatter plot on the right side, fitted with a fixed linear regression based on the same factors as in the effect plot. Bin number: trials of all sessions were averaged and subdivided into ten equally spaced units

A three-way interaction between task, stimulant intake, and age resulted in the best model fit (β = 0.86; CI = 0.36 – 1.35; see Fig. 6). Thus, over the course of a session subjects managed to generate more negative potentials, irrespective of the condition. The ability to produce negative SCP shifts in the activation and positive SCP shifts in the deactivation task (irrespective of time) was associated with both stimulant intake and age: being on regular stimulant medication and at higher age, or not being on

regular stimulant medication and at lower age. The inclusion of effects of the remaining factors summarized in Table 2 did not result in a better model fit.

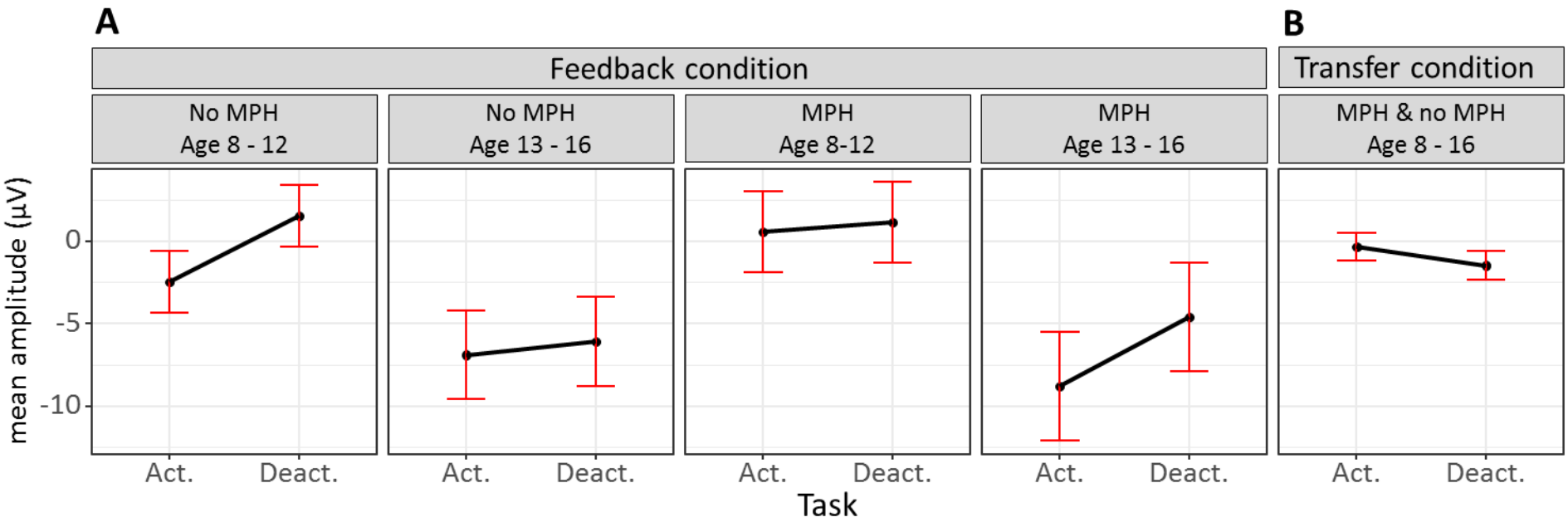

Figure 6. Visualization of within session NF-learning in the Feedback (**A**) and Transfer condition (**B**). The dependent variable is mean amplitude (µV) of baseline corrected trials. **A**: Interaction plot for the fixed effects task (deactivation versus activation), MPH and age. **B**: Effect plot for the fixed effect task (A*ct.: Activation; Deact.: Deactivation).* **A and B**: For visualization age is subdivided into two age classes (8-12 and 13-16 years), but preserved as a continuous variable in the original model.

The final model for within-session learning in the Transfer condition included subject as random intercept (τ_{00}= 13.320) and bin number as random slope (τ_{11}= 0.1231). As shown in Figure 7, a two-way interaction between bin number and condition resulted in the best model fit (β = 0.33; CI = 0.07 – 0.56; p= .011).

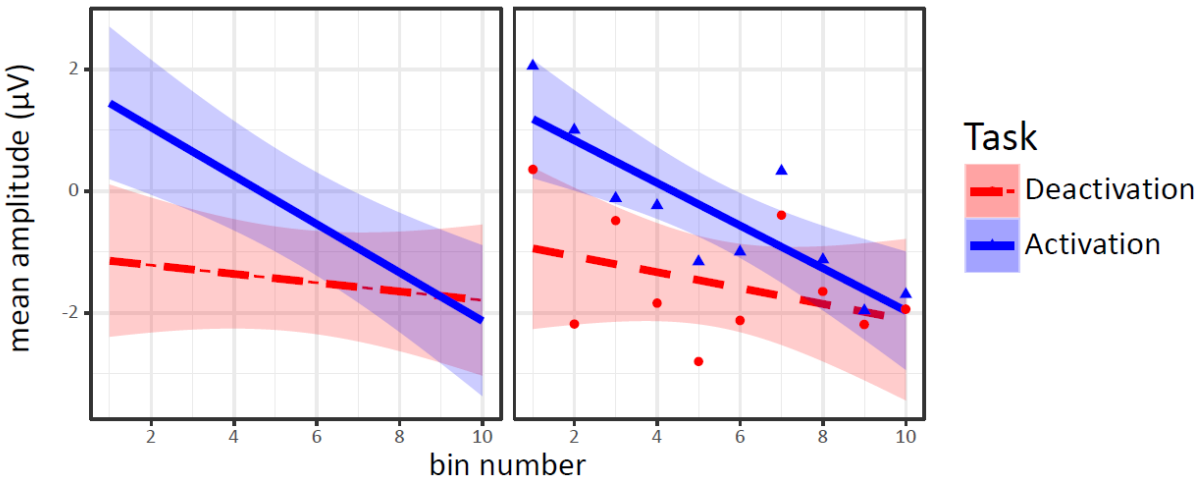

Figure 7. Visualization of within session-NF-learning in the Feedback condition. The dependent variable is the mean amplitude. The dependent variable n plot for the fixed effect bin number. For comparison between effect and raw data, see scatter plot on the right side, fitted with a fixed linear regression based on the same factors as in the effect plot. *Bin number:* trials of all sessions were averaged and subdivided into ten equally spaced units.

Age and MPH were included as additive fixed effects (age: β = -0.40; CI = -0.78– -0.00, p= .040; MPH: β = 1.71; CI = 0.03 – 3.40, p= .051; see Fig 8). Thus, being on constant stimulant medication resulted in an overall more positive mean amplitude.

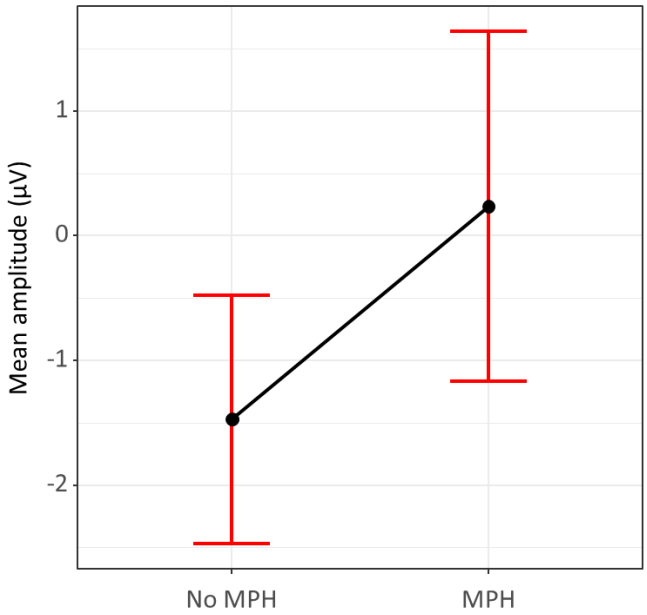

Figure 8. Effect plot of being on constant methylphenidate medication for within session learning in the Transfer condition. The dependent variable is mean amplitude (µV) of baseline corrected trials.

4.4.3 Additional analyses

We also analyzed whether NF learning was associated with the number of trials rejected due to artifacts by performing separate models for within and cross-session learning that included artifact rejection in the models. The mean artifact rate was 29.1% (±17%). The inclusion of the artifact rate did not yield a significantly better model fit for either condition.

To explore the number of subjects showing the desired learning slope in cross-session NF learning, models for both the Feedback and Transfer conditions were calculated separately and the subjects' random slopes were extracted to determine the individual learning performance for each task. Successful NF learning was defined by a negative slope in the activation task or a positive slope in the deactivation task. Subjects presenting both a positive slope in the deactivation task and a negative slope in the

activation task were labelled ‘successful regulators’. In the Feedback condition, 20 learners in the activation task, 23 learners in the deactivation task, and ten subjects were classified as successful regulators. In the Transfer condition, 23 subjects were classified as learners in the activation task, 23 as learners in the deactivation task, and eight as successful regulators.

4.5 Discussion

In this study, we analyzed differential effects on SCP -NF learning by considering within-session and cross-session analysis and direct feedback (Feedback condition) vs. delayed feedback (Transfer condition) using a mixed-models approach (Baayen et al., 2008).

4.5.1 Cross-session learning

In the Feedback condition, the desired NF learning pattern—a positive slope in the deactivation task and a negative slope in the activation task—was more pronounced with increasing age and regular intake of stimulant medication. In contrast, being younger and not taking stimulants was not associated with changes in NF learning across sessions, although the generation of potential shifts was still in the desired direction (mean amplitude in the activation task more negative than in the deactivation task). As being younger and not medicated was associated with a good differentiation between activation and deactivation right from training onset, one could argue that there could be little space for improvement. The NF literature offers little help in interpreting these seemingly inconsistent findings on the impact of medication and age, as studies allowing stimulants did not include these factors as covariates for NF

learning (Bakhshayesh et al., 2011; Drechsler et al., 2007; Gevensleben, Kleemeyer, et al., 2014b; Liechti et al., 2012; U. Strehl et al., 2006a; Vollebregt et al., 2014). It appears that maturation effects interacted here with the effects of stimulants. MPH intake might be confounded with other variables we did not take into account, such as the fact that children on medication still had to present clinically relevant symptoms to be included in the study. Therefore, one might speculate that a potential beneficial effect of stimulants on NF learning was not strong enough in younger children to compensate for ADHD severity. Similarly, age constitutes a proxy variable for developmental processes rather than explaining them (L. (Hrsg.). Oerter, R. & Montada, 1995; Trautner, 1991).

NF learning in the Transfer condition appeared to be especially challenging, as can be inferred from the relatively small potential shifts compared to the Feedback condition. Longer stimulant intake duration was associated with more pronounced NF learning in the deactivation task than in the activation task, which is difficult to explain. Simple fatigue cannot be the main reason, as this would produce a positive learning slope in both conditions. Increasing age was associated with a more negative mean amplitude irrespective of session number, which was probably related to larger proportions of fast frequencies as a function of age (Clarke et al., 2001; Matousek & Petersén, 1983; Matthis et al., 1981). Very few NF studies examining NF learning in ADHD have included adolescents (Van Doren et al., 2016). Thus, further research is needed to replicate our findings here too.

In the literature, findings on chronic MPH effects are very inconsistent. Despite reports on the beneficial and normalizing effects of long-term MPH medication (e.g. (Liotti et al., 2007)), other studies provide support for adaptation effects after chronic medication (Fusar-Poli, Rubia, Rossi, Sartori, & Balottin, 2012; Jensen et al., 2007) or even

showing no superiority of MPH over behavioral therapy after three years of constant MPH medication (Jensen et al., 2007).

4.5.2 Within-session learning

For within-session NF learning in the Feedback condition, subjects mainly generated more negative potential shifts throughout the session irrespective of condition, whereas in the Transfer condition, NF learning depended on the task: NF learning was especially evident in the activation task. At the beginning of the training, subjects did the opposite of what they were supposed to do (generating more positive potentials in the activation task and more negative potentials in the deactivation task). This finding was also reported by Strehl et al. (2007) for cross-session learning in the Transfer condition, with a normalization towards the desired polarities in the last training sessions.

We also considered whether double training sessions would be too exhausting for subjects with ADHD due to their limited attentional resources. However, our data do not show a pronounced positive increase in mean amplitude across bins, which would be expected with increased fatigue. For within-session NF learning in the Feedback condition, the ability to differentiate between activation and deactivation was especially prominent for being younger and not taking stimulants or being older and taking stimulants. Currently, no study on SCP NF has reported results on within-session learning (but see for within-session analyses for frequency band NF in ADHD Bink, van Nieuwenhuizen, Popma, Bongers, & van Boxtel, 2014; Escolano et al., 2014; Hillard, El-Baz, Sears, Tasman, & Sokhadze, 2013b; Janssen, Bink, Weeda, et al., 2016). As learning to generate potential shifts without immediate feedback is thought to be a good indicator for regulation capacities outside the laboratory (see Drechsler

et al., 2007), one could argue that the generalization of the acquired skills had not yet fully developed and would have needed more training sessions. Higher IQ was associated with a more negative mean amplitude. In former studies, a higher IQ was found to be associated with a higher proportion of fast frequencies, which might be responsible for a more negative mean amplitude (Gevensleben et al., 2009a).

Clinical symptoms or severity rated by parents and teachers did not contribute to a better model fit. This cannot simply be attributed to a lack of learning, as a sizable proportion of participants seemed to have learned to change brain activity in the desired direction. The artifact reduction rate has been shown in previous studies to improve over time as a possible non-specific effect of the treatment as children learn to sit still (Liechti et al., 2013), but whether this reduction in artifacts is related to NF learning has rarely been examined. In the present study, inclusion of the artifact reduction did not result in a better model fit, which suggests that even though artifact reduction took place across sessions, it was not related to NF learning within or across sessions. Gender has rarely been included as predictor due to the common overproportion of males. As almost 50% of our participants were females, we could test for possible gender differences in NF learning. However, including gender did not yield a better model fit. It was not surprising that setting was not associated with NF learning, as NF learning should not be affected by the training environment; however, differential setting effects on NF learning have never been tested before directly, so this needed to be shown. The inter-session interval has rarely been examined as an effect on NF learning or clinical improvement (Ros et al., 2009; Schabus et al., 2014). It did not yield a better model fit here, but this might also be attributable to small variations in the time schedule.

The analysis clearly shows interactions between such subjects' characteristics as age, stimulant intake, and session/bin number. Stimulant intake does not appear per se to be an impediment to NF learning; indeed, it seems to facilitate learning, at least in older children. A more systematic study with randomization of children on and off medication would be needed to analyze this association and replicate our findings. We expected our multilevel modeling approach would achieve a more realistic mapping of NF learning in ADHD than other statistical models, such as multivariate analysis. One limitation may be the lack of current consensus whether and if so to which degree it is possible to rely on p-values in mixed modeling (Baayen et al., 2008). Conversely, a major advantage of this statistical approach is that independence amongst observations is not a necessary pre-condition. The model presented here is also an approach to analysis that might be beneficial for future studies on the relation between NF learning and clinical outcome.

5 Paper 3 - Multilevel modeling of training specificity of neurofeedback in children and adolescents with Attention-Deficit/Hyperactivity Disorder – relating clinical symptoms to training performance

Agnieszka Zuberer[1], Franziska Minder[1], Daniel Brandeis[1-4], Renate Drechsler[1]

1 Department of Child and Adolescent Psychiatry, University of Zurich, Zurich, Switzerland
2 Neuroscience Center Zurich, University of Zurich and ETH Zurich, Switzerland
3 Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim/ Heidelberg University, Mannheim, Germany
4 Medical Faculty Mannheim/ Heidelberg University, Mannheim, Germany

5.1 Abstract

Background

In recent years, neurofeedback (NF) has gained increasing popularity as a method for training children and adults with ADHD. Analyses of NF training efficacy usually focus on clinical pre-post improvements but ignore specificity, that is, whether measures deployed during the training period might result to clinical improvement. It has been argued that tonic EEG changes are a more specific measure for training specificity in neurofeedback than training performance. However, tonic EEG changes have been a minor issue in NF research and have usually been analyzed with respect to pre-post treatment resting EEG without examining the carry-over effect of tonic changes throughout the whole training course.

Methods

Forty-four children and adolescents with ADHD underwent neurofeedback training of slow cortical potentials. Parent-rated clinical symptoms were assessed using the Conners-3 scales and the Behavior Rating Inventory of Executive Function (BRIEF) at three time points: 3 months before the start of training (T1), immediately before training (T2), and after the completion of training (T3). We deployed linear mixed-effect modeling to a) examine alpha power increments with respect to training session and within-session time (pre-post session resting baseline) and b) NF training performance (here termed NF learning) across sessions. Subject-related factors such as medication and age were included in the analysis. To analyze the relationship between clinical improvement across the three time points and EEG learning, we deployed a linear mixed-effect modeling approach to predict the change in clinical scores across time points with respect to either single-subject upper alpha slopes or NF learning slopes across training sessions.

Results

Parents' ratings of clinical symptom reduction from T2 to T3 on the Conners-3 Inattention DSM-IV Index and the BRIEF Behavioral Regulation (BRI) and Metacognition (MI) indices were associated with upper alpha increments across training sessions, whereas no such association could be reported for NF learning.

Discussion

Though not conclusive, our results suggest that tonic EEG measures might be a more reliable measure of treatment specificity than phasic measures.

5.2 Introduction

The focus of most neurofeedback training (NF) studies in ADHD has been on clinical outcome effects, assessed by pre-post changes in ratings of ADHD symptoms; training specificity has rarely been examined. One way to measure training specificity is to analyze whether the electrophysiological learning process across training sessions (which in this paper will be called NF learning) is related to clinical changes before and after training.

5.2.1 Association between clinical outcome and NF learning

A relatively small proportion of the considerable literature on clinical improvement and NF reports results on the association between NF learning and clinical outcome. Several studies have reported clinical outcome measures to be positively associated with NF learning (Drechsler et al., 2007; Gevensleben et al., 2010; Lubar et al., 1995; Strehl et al., 2006b), whereas other studies have not found any relationship (Escolano et al., 2014; Janssen, Bink, Geladé, et al., 2016; Liechti et al., 2012). ADHD symptoms rated by parents have been found to improve with learning (see Kropotov et al., 2005, for frequency band NF (FR-NF) and Gevensleben, Kleemeyer, et al., 2014a, for neurofeedback of slow cortical potentials (SCP), while other studies did not find any association between NF learning and clinical improvement based on parent and teacher ratings (Escolano et al., 2014; Janssen, Bink, Weeda, et al., 2016). Some studies dichotomized NF learning into good and poor learning by median split and found clinical improvements for parents' ratings to be positively associated with SCP-NF learning (Drechsler et al., 2007; Strehl et al., 2006a). In FR-NF, good learning was associated with improvements in attentional tests (Lubar, 1995) and Go/NoGo response time and variability (Kropotov et al., 2005).

5.2.2 Association between clinical outcome and resting baselines

Gruzelier (2014) argues that the original goal of neurofeedback is to modify tonic EEG and that NF learning should therefore be manifested in the progressive change of the pre-session EEG resting baselines of successive sessions. Hanslmayr et al. (2005) argue that neurofeedback is capable of inducing tonic EEG changes, at least in alpha power, in healthy adults rather than phasic, event-related changes. However, baseline increments across sessions have barely been examined in NF studies in ADHD, and the few studies that have analyzed them did not find any significant baseline increments (Janssen, Bink, Weeda, et al., 2016; Kropotov et al., 2005; Liechti et al., 2012). The relation between clinical outcome and changes in pre-post training baseline measures have been analyzed in several studies (Doehnert et al., 2008; Escolano et al., 2014; Gevensleben et al., 2009a; Janssen, Bink, Weeda, et al., 2016; Liechti et al., 2013). Some studies reported no association between pre-post changes in resting baselines and clinical outcome (Janssen, Bink, Weeda, et al., 2016). However, Doehnert et al. (2008) found theta/beta-ratio reductions and alpha increases to be correlated with reductions on the Conners Parent Rating Scale for hyperactivity/Impulsivity, and Gevensleben (2009) found an increase in alpha power to be related to improvements in hyperactivity/impulsivity. It has been common practice to subdivide alpha into two sub-bands due to their different functional associations. In healthy adults, enhanced upper alpha has been related to better cognitive performance (Doppelmayr, Klimesch, Hoedlmoser, Sauseng, & Gruber, 2005; Hanslmayr et al., 2005; Klimesch, Doppelmayr, & Hanslmayr, 2006). Neurofeedback studies employing a protocol to enhance upper alpha have found increases in upper alpha to improve cognitive performance in healthy adults (Hanslmayr et al., 2005; Zoefel et al., 2011). However, a study on upper alpha neurofeedback in ADHD found no association between resting EEG and behavioral measures (Escolano et al., 2014). Thus, the

present study addresses a dearth of research on possible associations between upper alpha and clinical outcome.

Clinical outcome is usually measured by the difference between clinical measures before and after training (Strehl et al., 2006; Drechsler et al., 2007). However, to account for possible waiting-time effects and statistical problems such as regression to the mean, we expect to obtain more reliable results from ratings at three time points. The goal of this study was to examine how the change of clinical measures across three time points might be related to a) the change in resting-baseline upper alpha power measured before and after each training session and b) NF learning in both the feedback and transfer conditions. By taking into account intra- and inter-subject variability of learning and clinical measures, a multilevel approach with linear mixed-effects modelling can provide a more complete picture of this association than basic multivariate analysis.

5.3 Methods

5.3.1 Participants

Children and adolescents with ADHD were recruited at schools, in outpatient clinics, by referral of clinicians, and by advertisement in parent self-aid groups for children with ADHD. See Table 1 for group characteristics.

Table 1. Description of the sample

N (total)	44
Male/female (N)	25/19
Setting school/clinic (N)	21/23
Age (years)	11.3 ± 2.1
Methylphenidate (with/without)	14/20
Dose (mg)	23.0 ± 12.1
Duration intake (years)	2.6 ± 2.6
IQ	108.8 ± 15.3

A precondition for participation was informed written consent signed by parents and the child. Approval for the study was given by the local ethics committee.

To be included in the study, subjects aged between 8.5 and 16.5 years had to present clinically relevant scores in the German versions of the Conners-3 Parent and Teacher Rating Scales (Lidzba, , Christiansen, H., Drechsler, 2013), according to DSM-IV ADHD indices T-scores (one of two ADHD DSM- IV indices T-value ≥ 65, the other T ≥ 60 according to both teacher and parent ratings for children of the combined subtype; ADHD DSM-IV Inattention Index T ≥ 65 in one and T ≥ 60 the other, rating for the inattentive subtype). For children taking methylphenidate (MPH), the same criteria had to be fulfilled, and ratings had to reflect behavior while on medication. In addition, the MPH dose had to be kept stable over the full treatment time (including three months

before the first measurement). Exclusion criteria were an estimated IQ < 80 (short form of the German WISC-IV; (Waldmann, 2008)), being treated with atomoxetine, neuroleptic, or any other psychoactive drug, severe comorbidities or other psychiatric disorders, and neurological disorders. Further exclusion criteria included insufficient knowledge of the German language, as participants and parents had to understand instructions and questionnaires, previous experience with NF (more than 4 lessons), and participation in or planning to start a treatment which might constitute a confound for training effects.

5.3.2 NF learning

For NF learning in the feedback condition, a three-way interaction between session number, condition (feedback or. transfer) and stimulant intake (MPH) yielded the best model fit ($\beta = 0.36$; CI = 0.17 – 0.56; p=.001). Thus, for the desired NF learning pattern, namely a steeper slope, was more pronounced when being older (within the age range of 8.5 – 16.5 years) and being on constant stimulant medication.

In the transfer condition, the best model fit was attained by a two-way interaction between session and MPH ($\beta = 0.42$; CI = 0.04 – 0.79; p=.034). Thus, being on constant stimulant medication was associated with better NF learning. Overall, the generation of potential shifts was smaller than in the feedback condition (see Table S2 in the Appendix and Table S4 for complete results).

5.3.3 Study design

Comorbid clinical conditions were screened using the Development and Well-Being Assessment (DAWBA) once with parents (Goodman et al., 2000). Parents' and

teachers' ratings of the children's behavior were completed at three time points: three months before training onset (screening = T1), directly before training onset (T2), and directly after training end (T3). We used the German adaptation of the Conners-3 Parent Rating Scale (Lidzba, Christiansen, Drechsler, 2013) for the measurement of ADHD symptoms and the German version of the Behavior Rating Inventory of Executive Function (Gioia, Isquith, Guy, & Kenworthy, 2000), parents' version, for the evaluation of executive function (EF) impairments in everyday life. This study was part of a more comprehensive project that included another treatment group and additional assessments.

As Table 2 shows, clinical improvements were found for all four measures from T2-T3, but only Conners-3 DSM-IV Inattention and Hyperactivity/Impulsivity Index scores improved from T1-T2, while the BRIEF Behavioral Regulation (BRI) and Metacognition indices remained stable (see Table S1 in Appendix B for detailed results -). A detailed discussion of these changes is to be found in Minder et al. (in prep.).

Table 2

Parents' ratings of clinical severity across three time points.

Conners-3 Parents DSM-IV indices (T-scores)	T1	T2	T3	T1 vs. T2	T3 vs. T2
Inattention	68.5 ± 5.1	65.0 ± 7.8	62.4 ± 8.2	T1>T2	T3<T2
Hyperactivity	65.3 ± 7.7	61.9 ± 8.8	59.1 ± 7.9	T1>T2	T3<T2
BRIEF Parents indices (T-scores)					
Metacognition	68.0 ± 12.0	66.7 ± 11.3	63.3 ± 9.9	T1=T2	T3<T2
Behavioral Regulation	62.0 ± 12.5	60.5 ± 12.3	55.0 ± 9.8	T1=T2	T3<T2

Note. Time points: three months before training onset (T1), directly before training onset (T2) and directly after training (T3). Direction of changes in clinical scores T1 vs. T2 and T3 vs. T2 is described in the last two columns. For detailed results of the mixed models analyses see supplement table 1.

Training was performed in outpatient clinics of the Department of Child and Adolescent Psychiatry and Psychotherapy of the University of Zürich (clinical setting, N=23) or at the subjects' school in a separate room (school setting, N=21). All training sessions were provided as individual therapy. One training program encompassed 30 sessions, usually administered as double sessions (two training sessions of 60 min. with a break in between). A complete training program consisted of 15 double sessions administered over 10 to 12 weeks: children who trained in the clinical setting started training as a 2-week vacation course with double training sessions daily (five double sessions per week; see Figure 1) followed by weekly double sessions over at least 5 weeks.

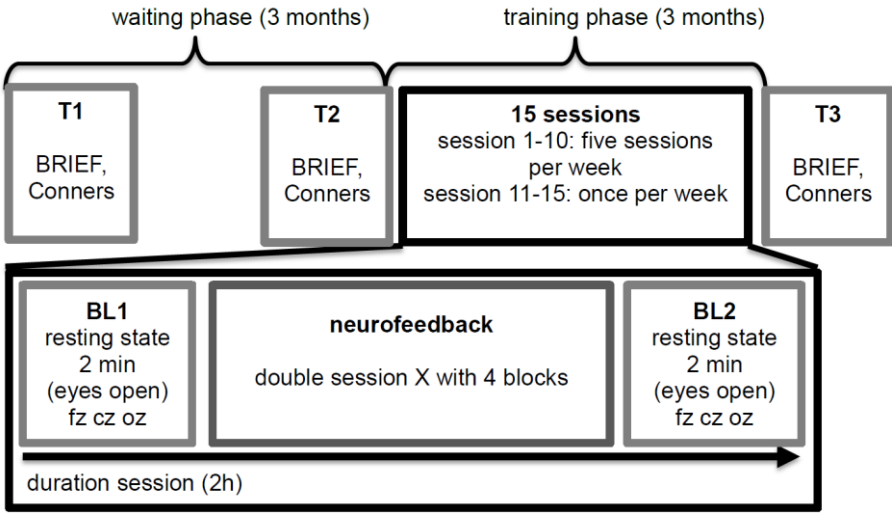

Figure 1. Design of the study and description of a training session

Training breaks of up to 10 days were allowed during the last training phase. Training in the school setting started with three sessions per week for the first two weeks. After that phase, seven weekly sessions followed. The different schedules in both settings were unavoidable; children in the clinical setting had to be trained in their leisure time and had only one afternoon off per week during the school period.

5.3.4 Description of the NF training

The NF training was provided from a commercially available mobile training device (Theraprax; NeuroConn GmbH) for SCP training, which also permitted Q-EEG recording. During a training session, the subject was seated in a comfortable chair in front of a computer monitor. The training was presented as a computer game in which the child had to steer a stimulus (e.g. fish, airplane) above or below a horizontal line while the stimulus moved from the left to the right side of the screen. The orientation of a triangle in the center of the screen indicated the task (activation or deactivation). The subject was instructed to generate either more positive potential shifts by a red upwards-pointing triangle or more negative potential sifts by a blue downward-pointing triangle. The change in cortical arousal was fed back by the target stimulus, which moved upwards or downwards proportionally to the degree of cortical excitability. One SCP trial lasted 8 seconds and consisted of three phases: a baseline phase (seconds 0–2), an active phase (8s), and a reinforcement phase (2s). The SCP in the active phase was baseline corrected.

Good performance was rewarded by a point at the end of the trial in the form of a sun. Depending on the condition, subjects had to modulate the NF parameter with or without the aid of feedback stimulus (feedback and transfer conditions). The proportion of activation and deactivation trials was 50%: 50%. All four types of trials appeared in randomized order. One double session consisted of four blocks, each containing forty trials. The percentage of transfer trials increased with session and block number (percentage of transfer trials per double session: double sessions 1-2: 20/ 20/ 20/ 20, double sessions 3-5: 20/ 20/ 20/ 40, double sessions 6-8: 20/ 20/ 40/ 40, double sessions 9-13: 20/ 40/ 40/ 50, and double sessions 14-15: 50/ 50/ 50/ 50).

Eyes-open resting EEG baselines were recorded before and after each training session (2 min). The subject was instructed to sit still and to look at a blue screen.

5.3.5 Montage and EEG recording

Both the NF training and resting EEG baseline recordings were performed at Cz, Fcz, and Oz electrodes and referenced to 2 mastoid electrodes shunted over a 10-kOhm resistance (Impedances < 20 kOhm; sampling rate was 512 Hz). Filtering of NF learning was performed from 0.01- 40Hz with a two-way least-squares FIR filter. For each 2.5s segment, a fast Fourier transformation was performed using the Welch method. Then, the mean of the power spectral density of upper alpha (10 - 12 Hz) was computed. The EEG amplifier (Theraprax Neuroconn©) used a low-pass filter of 40 Hz.

For the NF training, the slow cortical potentials (D.C. - 2 Hz) were processed from Cz channel for each sample point and displayed on the Thera Prax Trainer screen. There is a maximal time delay of 110 ms until the patient sees the feedback of the NF parameter. The vertical position of the feedback stimulus corresponds to the difference in amplitude with respect to the baseline of 2 seconds before the trial starts. Preprocessing was performed with MATLAB and EEGLAB.

5.3.6 Offline data preprocessing

NF learning and baseline recordings were filtered from 0.01- 40Hz with a two-way

least-squares FIR filter. For resting baselines, 2.5 sec segments of the continuous data recording were extracted. As regression-based artifact correction procedures did not yield reliable results, we applied a strict artifact removal procedure. After manual

artifact rejection, baseline-corrected trials for NF data or 2.5 sec segments for resting baseline data were rejected if their amplitudes exceeded ±100 mV or their gradients exceeded 50mv between two data points.

5.3.7 Statistical analysis

We employed a linear mixed-effects (LME) regression (Baayen, 2008; D. Bates, Mächler, Bolker, & Walker, 2015) to investigate the effects of neurofeedback training on clinical severity (dependent variable) with respect to three time points (T1: three months before training onset, T2: before training onset and T3: after training completion). Both T1 and T3 are referenced to T2. In the discussion, the time changes are termed T1-T2 and T2-T3. Four clinical measures were chosen as dependent variables to indicate clinical severity: the DSM-IV Inattention and Hyperactivity/Impulsivity indices (T-scores) of the Conners-3 Parent Rating Scale and the Behavioral Regulation (BRI) and Metacognition indices (T-scores) of the BRIEF parent rating scale. As a further independent variable, random slopes from separate mixed models analyzing NF learning or cross-session resting upper alpha (10-12.5 Hz) power change were included. The mixed model results for NF learning and upper alpha power can be seen in the supplement (S3 and S5 respectively). NF learning slopes and upper alpha power slopes were analyzed in separate models. In total, twelve models were tested (see table 3).

Table 3

Description of the analysis of the relationship between clinical outcome and NF-learning and Alpha-slopes.

		Clinical Severity (Dependent Variables)			
Random (Independent Variables)	Slopes	Inattention Index DSM-IV (Conners)	Hyperactivity / Impulsivity Index DSM-IV (Conners)	Metcognition Index (BRIEF)	Behavioural Regulation Index (BRIEF)
NF-Learning:					
	Feeback	n.s.	n.s.	n.s.	n.s.
	Transfer	n.s.	n.s.	n.s.	Model 2
Resting Upper Alpha Power		Model 1	n.s.	Model 4	Model 3

Note. in total 12 models were analyzed. Each model included one of the dependent and independent variables listed in the table above. Models where the inclusion of the independent varibale did not result in a significantly better model fit are labeled with n.s.. *Neurofeedback –Learning / Resting Upper Alpha Power*: Random slopes of a linear mixed effects model predicting Neurofeedback-learning across 15 sessions (*NF-learning*) or predicting change in resting upper alpha power across 15 sessions (*Upper Alpha Power*). *Conners*: Conners-3 Parent scale (T-scores). *BRIEF:* Behavior Rating Inventory of Executive Function parent scale (T-scores). *Feedback*: Type of neurofeedback condition, where feedback stimulus is provided. *Transfer*: Type of neurofeedback condition, where feedback is delayed.

For the analysis of NF learning, the mean differences of SCP amplitudes per session during both tasks (positivity/negativity) were computed for feedback condition and transfer condition separately (dependent variable). Thus, feedback NF learning and transfer NF learning were analyzed in different models.

For the analysis of upper alpha resting baselines, the mean power spectral density for each session was computed for each pre-session and post-session recording separately. For all model analyses, a random effect was included if there was a significant difference between the log-likelihood ratio ($p < 0.05$) and a difference in the AIC of at least 5.Higher-order interactions were kept in the model due to the principle of marginality (DebRoy & Bates, 2003). P-values were computed using a Kenward-Roger approximation. Confidence intervals were calculated with a bootstrapping

procedure. Table S1 in the Appendix B displays the effects considered for analysis. Statistical analysis was performed using the lme4 package (Bates et al., 2015).

5.4 Results

5.4.1 Association between clinical severity and EEG measures

The only models reported are those that showed a significant interaction with the learning measures described above. Interaction effects between these EEG measure slopes and the clinical indices (Conners-3 DSM-IV Inattention and Hyperactivity/Impulsivity parent indices; BRIEF Behavioral Regulation and Metacognition parent indices) across three time points are represented in Figures 2 to 4. Correlation between the random slope effects of session of the upper alpha power and the NF learning models (feedback condition) was not significant (Pearson r= -0.12; p=0.16).

Four models yielded a better model fit when including one learning measure. Results of mixed models analyzing the relationship between EEG measures and clinical scores are summarized in Table 4.

Table 4

Results for linear mixed effects models with dependent variables in the titles with the independent variables of interest in brackets.

	Inattention DSM-IV index (uA slope)			Behavioural regulation index (NF - TR)			Behavioural regulation index (uA slope)			Metacognition (uA slope)		
	B	*CI*	*p*	*B*	*CI*	*p*	*B*	*CI*	*p*	*B*	*CI*	*p*
Fixed Parts												
Intercept	64.37	62.22 – 66.52	**<.001**	60.19	56.81 – 63.57	**<.001**	59.75	56.21 – 63.30	**<.001**	66.12	62.81 – 69.43	**<.001**
T1	4.04	2.11 – 5.98	**<.001**	1.86	-1.29 – 5.01	.251	1.72	-1.57 – 5.01	.308	1.81	-1.40 – 5.01	.272
T3	-1.82	-3.78 – 0.15	.074	-5.68	-8.88 – -2.47	**.001**	-4.66	-8.02 – -1.30	**.008**	-2.76	-6.03 – 0.50	.101
slope uA	115.78	-73.45 – 305.01	.234				314.89	-0.13 – 629.91	.053	323.34	33.26 – 613.42	**.031**
T1 vs. T2 : slope uA	-322.49	-491.69 – -153.30	**<.001**				-55.70	-367.26 – 255.86	.727	54.97	-242.42 – 352.36	.718
T3 vs. T2 : slope uA	-202.88	-378.60 – -27.15	**.026**				-391.70	-697.00 – -86.40	**.013**	-457.05	-747.88 – -166.23	**.003**
slope NF (TR)				-12.08	-22.52 – -1.63	**.026**						
Random Parts												
σ^2		20.221			52.435			55.334			53.616	
$\tau_{00,\ subject}$		30.831			73.171			78.671			65.621	

Only models are reported where the EEG-measures (NF-learning and upper alpha slope) improved the model-fit. *T* : time point of clinical severity: 3 months (T1), directly before training onset (T2), and after training completion (T3). Reference is T2. *Inattention* and *Hyper-activity/Impulsivity DSM-IV index :* rated by the Conners-3 Parent scale (T-scores). *Behavioural regulation* and *Metacognition index* : rated by BRIEF parent scale (T-scores). *NF :* Neurofeedback. *NF slope / uA slope :* Random slopes of a linear mixed effects model predicting NF across 15 sessions (NF slope) or predicting change in resting upper alpha power (eyes open) across 15 sessions. *TR:* Type of neurofeedback task, where feedback is delayed (Transfer). $t_{00,\ subject}$: between-subject variance.

As Figure 2 shows, an increase in inattention symptoms from T1 to T2 was associated with a steeper alpha slope (β = -322.49; CI =-491.69 – -153.30; p= <.001), while showing a decrease in symptom severity from T2 to T3 was associated with a shallower upper alpha slope (β = -202.88; CI =-378.60 – -27.15; p= .026). Change in Hyperactivity/Impulsivity Index scores across the three time points was associated with neither NF learning nor cross-session alpha changes; in consequence, it is not represented here.

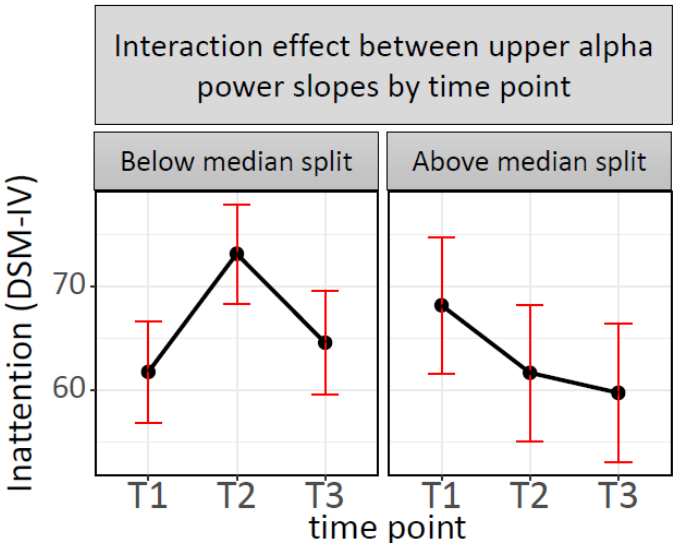

Figure 2. Effects for the dependent variable Inattention. Interaction between time points and upper alpha power slopes. Upper alpha power slopes: random slopes of a linear mixed effect model predicting change in resting upper alpha power (pre-

BRI T-scores decreased from T2 to T3 (β = -5.86; CI =-8.88 – -2.47; p= .001), whereas no significant change was found for T1 to T2 (β = 1.86; CI =-1.29 – 5.01; p= .251) (Figure 3A). As seen in figure 3B, a steeper NF-learning slope in the Transfer condition was associated with more severe BRI scores (β = -12.08; CI =-22.52 – -1.63; p= .026). However, there was no interaction between NF slope and BRI changes across the three time points.

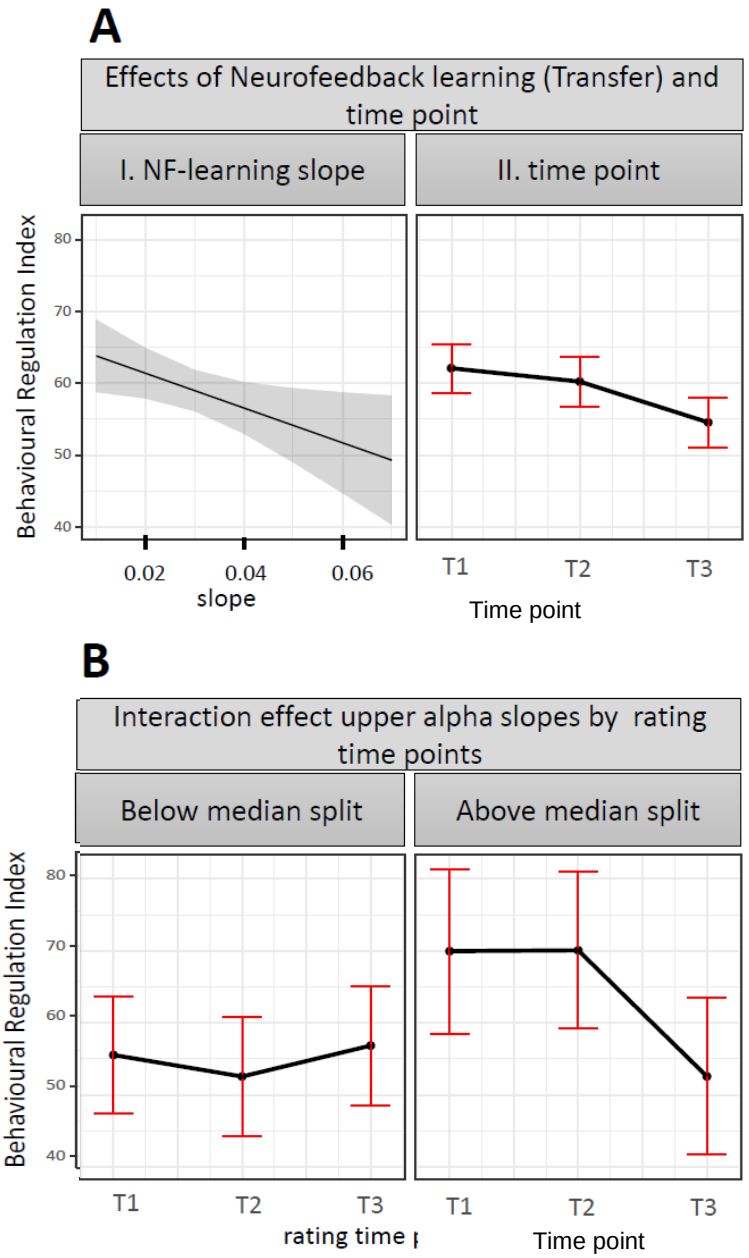

Figure 3. Effects plot for the dependent variable Behavioral Regulation Index (BRI). **(A)** I. Effect of NF learning slopes (Transfer) on BRI; II. Effect of time point on BRI. No interaction between NF learning slopes and time point on BRI. **(B)** Interaction between time point and upper alpha slopes of the resting baselines. *NF-learning slopes*: random slopes of a linear mixed effects model predicting NF-learning across 15 sessions. *Upper alpha slopes*: random slopes of a linear mixed effect model predicting change in resting upper alpha power across 15 sessions. *Time point*: *T1:* three months before training onset, *T2*: before training onset, *T3:* after training completion. *Upper alpha slopes is* factorized by median split for visualization but preserved as a continuous variable in the original model. BRIEF Behavioral Regulation Index parent scale (T-scores).

A steeper cross-session increase in upper alpha power (see Figure 3B) was associated with a decrease in BRI T-scores for T2 to T3 (β = -391.70; CI = -697.00 – -86.40; p= .013), whereas no significant change was found for T1 to T2 (β = -55.70; CI = -367.26 – 255.86; p= .727). A cross-session increase in alpha power (Figure 4, above median split) was associated with a decrease in Metacognition Index scores for T2 to T3 (β = -457.05; CI = -747.88 – -166.23; p= .003), whereas there was no pronounced change for T1 to T2 (T1 to T2: β = 54.97; CI =, β = -457.20; CI = -242.42 – 352.36 ; p= .718).

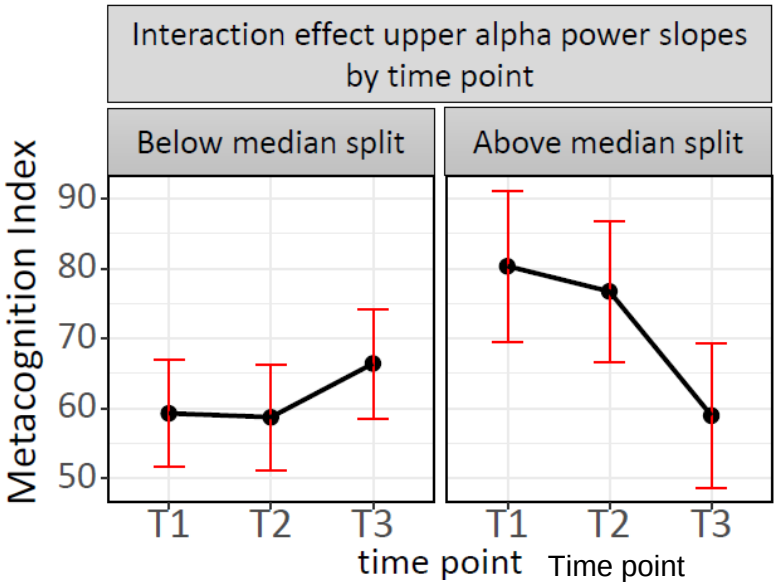

Figure 4. Effects plot for the dependent variable Metacognition Index. Interaction between rating time point and upper alpha power slopes of the resting baselines. *Alpha power slopes*: Random slopes of a linear mixed effect model predicting change in alpha power resting BL across 15 sessions. *Upper alpha power slopes* are factorized by median split for visualization but preserved as a continuous variable in the original model. *Time point*: T1: Three months before training onset, T2: before training onset, T3: after training completion. BRIEF Metacognition index parent scale (T-scores).

The NF slopes of both feedback and transfer conditions were unrelated to the Metacognition Index score and thus are not depicted here.

5.6 Discussion

In this paper, we investigated the degree to which cross-session EEG learning parameters (the NF slope in the feedback condition, the NF slope in the transfer condition, and the upper alpha power slope) were related to clinical outcome gains measured by parents' ratings of ADHD symptoms and executive functions (EFs). The rationale behind this approach was to develop a valid measure for treatment specificity: a good clinical outcome should be reflected in improved regulation of the trained EEG parameter. If a clinical improvement is not associated with training performance, either effects unrelated to the electrophysiological aspects of the NF training must be responsible, or the measures are not valid.

The decrease in parent-rated Inattention (Conners-3 DSM-IV Index) scores across the three time points (T1 to T2, T2 to T3) was not associated with NF learning. Changes in inattention symptoms were, however, associated with the upper alpha slope: a stronger decrease in inattention scores from T2 to T3 (training period) was associated with a steeper upper alpha slope. However, a decrease of inattention scores from T1 to T2 (waiting period) was associated with a less steep or even negative upper alpha slope. Although hyperactivity/impulsivity (Conners-3 DSM-IV Index) scores significantly decreased across the three time points, changes were neither associated with NF learning nor with cross-session upper alpha change.

For EF impairment as measured by the BRIEF parent rating scale, less severe Behavioral Regulation (BRI) index symptoms were associated with better NF learning (i.e., a steeper slope) in the transfer condition. However, this finding was not related to the time points, although behavioral regulation symptoms improved across the training (T2 to T3). No association was found between the Metacognition Index and NF

learning. Improvements in both Metacognition and BRI scores across the training period (T2 to T3) were associated with a steeper upper alpha slope.

The fact that NF learning was not related to clinical outcome contradicts assumptions that clinical improvements are a result of electrophysiological learning processes. However, the finding that reductions in Inattention, Behavioral Regulation and Metacognition index symptoms from T2 toT3 were associated with stronger alpha power increments across sessions indicates that certain training-related effects occurred. An association between alpha power and clinical symptoms or executive functions finds support in previous research, where reductions in alpha and beta power have been associated with learning disabilities in school age children (for review see Klimesch et al., 1999). In two studies of healthy adults, alpha power was increased during retention phases (Jensen et al., 2002; Klimesch et al., 1999), possibly representing bottom-down processes inhibiting interfering input. Klimesch and colleagues (1999) argue that these inhibition processes might be subsumed by central executive processes. These findings might explain why improvements in both BRIEF indices after training completion were related to alpha enhancements across training sessions. However, this should not lead to the conclusion that a positive alpha slope is an index of NF training specificity. Alpha increments could also have occurred as a consequence of non-specific training effects, such as attentional demands during NF training. We found no systematic association between changes in upper alpha power and NF learning, which is also an argument against training specificity. However, NF learning did not take place uniformly across subjects; it interacted with age and stimulants intake. Conversely, upper alpha changes across training sessions did not interact with other effects but were only associated with additive effects such as age, electrode site, and time (pre-session to post-session). Thus, systematic interactions

between clinical changes and upper alpha slope may be easier to find than with NF learning.

The divergent results with BRIEF and Conners-3 scales for the waiting period (T1-T2) raise another question. A stronger improvement in inattention on the Conners-3 scale from T1 to T2 was associated with a steeper upper alpha slope, but this is counterintuitive, as improvements during waiting time are by definition unrelated to training effects. However, irrespective of any association of symptoms with EEG measures, Conners-3 indices, but not BRIEF indices, showed significant changes throughout the waiting phase (T1 to T2). One possible reason for this discrepancy might be that parents had been biased by rating ADHD symptoms on Conners-3 scales at T1, which was used for screening before entering the study. Another explanation for this unexpected waiting time effect is that greater symptom reductions in relation to upper alpha enhancements across training sessions were associated with more severe clinical symptoms at T1. Thus, more severe clinical symptoms at T1, which were related to reductions in upper alpha power, arguably allowed more space for improvement. However, the inclusion of the subjects' random intercepts of upper alpha power into the model did not yield a better model fit.

5.6.1 Conclusion

The evidence presented in this study for specific effects of NF training on clinical outcome is relatively poor, as we only found effects for upper alpha baseline increases, but not for NF learning. That NF learning might be a less specific measure than tonic EEG has been argued by Gruzelier (2014), who claimed the carry-over effect on resting baselines to be “the more enduring goal for feedback learning”. However, as our training parameter was SCPs and not frequency bands, direct conclusions cannot be drawn.

Thus, further research is needed to replicate our findings for frequency band NF and to investigate the association of other frequency bands with clinical outcome, also with respect to a proper control group. This is one of the few studies using a multilevel approach to investigate the association between clinical outcome and resting-baseline increments across sessions in ADHD. As a consequence, this paper also constitutes an example of a methodological approach that can possibly yield more valid results by accounting for inter- and inter-subject variability than statistical methods such as analyses of variance.

6 General discussion

6.1 Measuring NF-treatment specificity

As discussed in paper 1, one major issue in measuring NF-treatment efficacy in ADHD is to relate NF-learning to the clinical outcome. However, most studies examining these issues have employed basic statistical methods and did not include subject and context-related factors. Thus, many questions on training efficacy and specificity remain unresolved.

6.2 Prediction of NF-learning

In paper two, we presented results on the prediction of NF-learning. The results show that NF-learning is dependent on subject related factors. As such, NF-learning across sessions was more advanced with increasing age and intake of stimulant and when performing neurofeedback trials where constant performance feedback was provided. In contrast, when subjects had to perform without the aid of a performance feedback, a stimulant effect remained, but learning was not associated with age. Within-session analysis showed that previous concerns about fatigue due to long training sessions (120 minutes) were unjustified, as subject became more alert within the course of a session. Contrary to our expectations, the inclusion of ADHD symptoms did not result in a better model fit. Similarly, the inclusion of inter-session-time interval (in days), dosage of stimulant medication, sex, setting, muscle artifact rate did not result in a better model fit. This is one of the rare studies employing a linear mixed effects approach, controlling for inter- and intra-subject variability. Thus, further research is needed to replicate our results. To sum up, the data showed, that NF-learning could not be predicted by ADHD symptoms or severity. Stimulant intake per se did not appear

to be an impediment to NF learning, but, at least in older children, seemed to facilitate learning.

6.3 Association between NF-learning and clinical outcome gains

In the third paper, we analyzed if, and if yes, to which extend changes in clinical severity across three rating time points (T1: three months before training onset, T2: directly before training onset, T3: directly after training end) can be explained by measures of NF-learning and resting EEG alpha-power across sessions. By this approach we assumed to tackle problems related to the use of only two time points, such as regression to the mean or placebo and developmental effects. For ratings on clinical severity we employed a) measures of inattention and hyperactivity/impulsivity (DSM-IV scores of the Conners-3 parent scale (Lidzba et al., 2013)) and b) metacognition and behavioral regulation indices of the BRIEF rating scale (Drechsler & Steinhausen, 2013). We found that a stronger decrease in Inattention scores from T2 to T3 (training period) was associated with less or no upper alpha enhancements, whereas an increase of Inattention scores from T1 to T2 (waiting period) was also associated with less or no upper alpha enhancements.

Less severe Behavioral Regulation (BRI) symptoms were associated with better NF-learning (i.e. a steeper slope) in the Transfer condition. However, this finding was not related to the rating time points. Nevertheless, behavioral regulation symptoms improved across the training (T2 to T3). We found no association between the Metacognition Index and NF learning. Improvements of both Metacognition and BRI scores across the training period (T2 to T3) were associated with stronger alpha enhancements. These findings raise questions towards the validity for the

measurement of NF-induced change with the Conners-3 scale and the validity of NF learning as a measure of training specificity.

6.4 Is absolute alpha power a pertinent electrophysiological measure to analyze treatment specificity?

Besides the previous discussion on how and if direct conclusions from the trained parameter in our NF training (namely SCPs) can be made about changes in frequency bands in the resting EEG, it t is important to ask, if frequency bands as we employed them, are a pertinent measure to analyze possible treatment specificity.

There is some evidence that fixed frequency bands are sensitive to error, not only due to age related variations in peak frequencies but also because a large variability across subjects of comparable age. For instance, a way to distinguish theta from alpha frequency is related to the fact that a frequency expected within the traditional alpha range desynchronizes during an active test condition as compared to a resting condition, whereas theta does not. This finding is a crucial marker to determine the individual alpha peak frequency (IAF). Specifically, the IAF is determined by those frequency points where the first derivative of the spectrum changed from positive to negative. By looking for peaks it is guaranteed to extract true peaks from the spectrum and not instead values at the boundary of the predefined alpha range. After peak extraction the expected alpha window of 5HZ is defined in reference to the individual alpha peak frequency. Several studies have shown superiority of this individually defined alpha frequency band over the traditional fixed alpha frequency band range (for discussion see Klimesch et al., 1999). Besides absolute power measures, the Alpha-Rhythm Dynamic Range has been found to be related to clinical symptoms, at

least in visuo-spatial neglect, after alpha-NF training (Ros, Michela, Bellman, Vuadens, & Saj, 2017). Further research is needed to see if the Alpha-Rhythm Dynamic Range is also a valid measure for NF treatment specificity in ADHD.

6.5 What are pertinent measures for clinical symptom improvement?

In the literature on ADHD interventions, the traditional view is that symptoms of ADHD are the behavioral consequences of cognitive deficits (see discussion in Coghill, Rhodes, & Matthews, 2007). Thus, it is hypothesized that when taking into account behavioral ratings, conclusion about cognitive functioning can be made. However, there is no or only weak scientific support for this assumption in the literature. In fact, for the administration of methylphenidate, studies failed to show any correlation between improvements in cognition and improvements in ADHD symptoms (Coghill et al., 2007). Thus, further research is needed to elucidate the relationship between clinical improvements and the inter-relation between both cognitive functioning and symptom ratings. Taking these findings together, our results with respect to symptom ratings need to be regarded with caution when deriving any conclusions about the improvement in the cognitive domain and thus better functioning in everyday life.

Furthermore, additional measures to symptom ratings through teachers/parents or cognitive testing might further contribute to measuring training efficiency. As such, systematic classroom observation procedures of behavior appear to be promising measurement instruments to measure stability/change in ADHD symptomology (Carboni, Roach, & Fredrick, 2013; Steiner, Frenette, Rene, Brennan, & Perrin, 2014).

6.6 Conclusion/Outlook

The two studies suggest that treatment specificity can be examined by employing learning measures of the actual training. However, the phenomenology of the association between NF learning and clinical outcome has to be further elucidated. In the first study it was suggested that age and stimulants intake may play a major role in NF-learning, but that NF-learning cannot be predicted by other subject-related factors or ADHD subtype. The finding of the second paper, namely that baseline alpha increments might be related to improvements in executive functions, but not in ADHD symptoms needs further investigation. Cross-validations with a new data set might yield additional results which cannot be answered with the present NF data alone.

7 References

Allison, B., & Neuper, C. (2010). Could Anyone Use a BCI? In *Brain-Computer Interfaces SE - 3* (pp. 35–54). http://doi.org/10.1007/978-1-84996-272-8_3

Arns, M., Conners, C., & Kraemer, H. (2013). A Decade of EEG Theta/Beta Ratio Research in ADHD: A Meta-Analysis. *Journal of Attention Disorders*, *17*(5), 374–383. http://doi.org/10.1177/1087054712460087

Arns, M., Drinkenburg, W., & Leon Kenemans, J. (2012). The effects of QEEG-informed neurofeedback in ADHD: an open-label pilot study. *Applied Psychophysiology and Biofeedback*, *37*(3), 171–80. http://doi.org/10.1007/s10484-012-9191-4

Arns, M., Heinrich, H., & Strehl, U. (2014). Evaluation of neurofeedback in ADHD: The long and winding road. *Biological Psychology*. http://doi.org/10.1016/j.biopsycho.2013.11.013

Baayen, R. H. (2008). Analyzing Linguistic Data: A Practical Introduction to Statistics Using R. Harald Baayen (2008). *Sociolinguistic Studies*, *2*(3), 353. http://doi.org/10.1558/sols.v2i3.471

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. http://doi.org/10.1016/j.jml.2007.12.005

Bakhshayesh, A. R., Hänsch, S., Wyschkon, A., Rezai, M. J., & Esser, G. (2011). Neurofeedback in ADHD: a single-blind randomized controlled trial. *European Child & Adolescent Psychiatry*, *20*(9), 481–91. http://doi.org/10.1007/s00787-011-0208-y

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. http://doi.org/10.1016/j.jml.2012.11.001

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. http://doi.org/10.1177/009286150103500418

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. http://doi.org/10.18637/jss.v067.i01

Bink, M., van Nieuwenhuizen, C., Popma, A., Bongers, I. L., & van Boxtel, G. J. M. (2014). Neurocognitive effects of neurofeedback in adolescents with ADHD: a randomized controlled trial. *The Journal of Clinical Psychiatry*, *75*(5), 535–42. http://doi.org/10.4088/JCP.13m08590

Birbaumer, N., Elbert, T., Canavan, G. M. A., & Rockstroh, B. (1990). Slow Potentials of the Cerebral Cortex and Behavior. *Physiological Reviews*, *70*(164), 1–41.

Cannon, R., Lubar, J., Congedo, M., Thornton, K., Towler, K., & Hutchens, T. (2007). The effects of neurofeedback training in the cognitive division of the anterior cingulate gyrus. *The International Journal of Neuroscience*, *117*(3), 337–57. http://doi.org/10.1080/00207450500514003

Carboni, J. a., Roach, A. T., & Fredrick, L. D. (2013). Impact of mindfulness training on the behavior of elementary students with attention-deficit/hyperactive disorder. *Research in Human Development*, *10*(3), 234–251. http://doi.org/10.1080/15427609.2013.818487

Caria, A., Veit, R., Sitaram, R., Lotze, M., Weiskopf, N., Grodd, W., & Birbaumer, N. (2007). Regulation of anterior insular cortex activity using real-time fMRI. *NeuroImage*, *35*(3), 1238–1246. http://doi.org/10.1016/j.neuroimage.2007.01.018

Cerquera, A., Arns, M., Buitrago, E., Gutiérrez, R., & Freund, J. (2012). Nonlinear dynamics measures applied to EEG recordings of patients with attention deficit/hyperactivity disorder: quantifying the effects of a neurofeedback treatment. *Conference Proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, *2012*, 1057–60. http://doi.org/10.1109/EMBC.2012.6346116

Cho, M. K., Jang, H. S., Jeong, S.-H., Jang, I.-S., Choi, B.-J., & Lee, M.-G. T. (2008). Alpha neurofeedback improves the maintaining ability of alpha activity. *Neuroreport*, *19*(3), 315–317.

Cinciripini, P. M. (1984). Discrimination of sensorimotor EEG (12-15 Hz) activity: a comparison of response, production, and no-feedback training conditions. *Psychophysiology*, *21*, 54–62.

Clarke, A. R., Barry, R. J., McCarthy, R., & Selikowitz, M. (2001). Age and sex effects in the EEG: Development of the normal child. *Clinical Neurophysiology*, *112*(5), 806–814. http://doi.org/10.1016/S1388-2457(01)00488-6

Clarke, A. R., Barry, R. J., McCarthy, R., Selikowitz, M., & Brown, C. R. (2002). EEG evidence for a new conceptualisation of attention deficit hyperactivity disorder. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *113*(7), 1036–44. http://doi.org/10.1016/S1388-2457(02)00115-3

Coghill, D. R., Rhodes, S. M., & Matthews, K. (2007). The Neuropsychological Effects of Chronic Methylphenidate on Drug-Naive Boys with Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry*, *62*(9), 954–962. http://doi.org/10.1016/j.biopsych.2006.12.030

Cortese, S., Ferrin, M., Brandeis, D., Holtmann, M., Aggensteiner, P., Daley, D., … Zuddas, A. (2016). Neurofeedback for Attention-Deficit/Hyperactivity Disorder: Meta-Analysis of Clinical and Neuropsychological Outcomes From Randomized Controlled Trials. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*(6), 444–455. http://doi.org/10.1016/j.jaac.2016.03.007

Daum, I., Rockstroh, B., Birbaumer, N., Elbert, T., Canavan, a, & Lutzenberger, W. (1993). Behavioural treatment of slow cortical potentials in intractable epilepsy: neuropsychological predictors of outcome. *Journal of Neurology, Neurosurgery, and Psychiatry*, *56*(1), 94–7. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1014773&tool=pmcentrez&rendertype=abstract

De Zambotti, M., Bianchin, M., Magazzini, L., Gnesato, G., & Angrilli, A. (2012). The efficacy of EEG neurofeedback aimed at enhancing sensory-motor rhythm theta ratio in healthy subjects. *Experimental Brain Research*, *221*(1), 69–74. http://doi.org/10.1007/s00221-012-3148-y

DebRoy, S., & Bates, D. M. (2003). Computational methods for single level linear mixed-effects models. *Department of Statistics, University of Wisconsin-Madison*. Retrieved from http://pages.cs.wisc.edu/~bates/reports/TR1073dist-a4.pdf%5Cnpapers2://publication/uuid/05C9257F-9BA7-4418-8B70-162D1C179636

Dempster, T., & Vernon, D. (2009). Identifying indices of learning for alpha neurofeedback training. *Applied Psychophysiology and Biofeedback*, *34*(4), 309–28. http://doi.org/10.1007/s10484-009-9112-3

Dickhaus, T., Sannelli, C., Müller, K.-R., Curio, G., & Blankertz, B. (2009). Predicting BCI performance to study BCI illiteracy. *BMC Neuroscience*, *10*(Suppl 1), P84. http://doi.org/10.1186/1471-2202-10-S1-P84

Doehnert, M., Brandeis, D., Straub, M., Steinhausen, H.-C., & Drechsler, R. (2008). Slow cortical potential neurofeedback in attention deficit hyperactivity disorder: is there neurophysiological evidence for specific effects? *Journal of Neural Transmission*, *115*(10), 1445–56. http://doi.org/10.1007/s00702-008-0104-x

Donfrancesco, R., Di Trani, M., Gregori, P., Auguanno, G., Melegari, M. G., Zaninotto, S., & Luby, J. (2013). Attention-deficit/hyperactivity disorder and alexithymia: A pilot study. *ADHD Attention Deficit and Hyperactivity Disorders*, *5*(4), 361–367. http://doi.org/10.1007/s12402-013-0115-9

Doppelmayr, M., Klimesch, W., Hoedlmoser, K., Sauseng, P., & Gruber, W. (2005). Intelligence related upper alpha desynchronization in a semantic memory task. *Brain Research Bulletin*, *66*(2), 171–177. http://doi.org/10.1016/j.brainresbull.2005.04.007

Drechsler, R., Straub, M., Doehnert, M., Heinrich, H., Steinhausen, H.-C., & Brandeis, D. (2007). Controlled evaluation of a neurofeedback training of slow cortical potentials in children with Attention Deficit/Hyperactivity Disorder (ADHD). *Behavioral and Brain Functions*, *3*(35), 35. http://doi.org/10.1186/1744-9081-3-35

Egner, T., & Gruzelier, J. H. (2001). Learned self-regulation of EEG frequency components affects attention and event-related brain potentials in humans. *Neuroreport*, *12*(18), 4155–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11742256

Egner, T., Strawson, E., & Gruzelier, J. H. (2002). EEG signature and phenomenology of alpha/theta neurofeedback training versus mock feedback. *Applied Psychophysiology Biofeedback*, *27*(4), 261–270. http://doi.org/10.1023/A:1021063416558

Enriquez-Geppert, S., Huster, R. J., Scharfenort, R., Mokom, Z. N., Vosskuhl, J., Figge, C., … Herrmann, C. S. (2013). The morphology of midcingulate cortex predicts frontal-midline theta neurofeedback success. *Frontiers in Human Neuroscience*, *7*(August), 453. http://doi.org/10.3389/fnhum.2013.00453

Enriquez-Geppert, S., Huster, R. J., Scharfenort, R., Mokom, Z. N., Zimmermann, J., & Herrmann, C. S. (2014). Modulation of frontal-midline theta by neurofeedback. *Biological Psychology*, *95*, 59–69. http://doi.org/10.1016/j.biopsycho.2013.02.019

Escolano, C., Aguilar, M., & Minguez, J. (2011). EEG-based upper alpha neurofeedback training improves working memory performance. *Conference Proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, *2011*, 2327–30. http://doi.org/10.1109/IEMBS.2011.6090651

Escolano, C., Navarro-gil, M., Garcia-campayo, J., Minguez, J., Escolano, C., Navarro-gil, M., … Minguez, J. (2014). The Effects of Individual Upper Alpha Neurofeedback in ADHD : An Open-Label Pilot Study, 1–10.

Faraone, S. V., Biederman, J., & Mick, E. (2006). The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies. *Psychological Medicine*, *36*(2), 159–65. http://doi.org/10.1017/S003329170500471X

Fayyad, J., De Graaf, R., Kessler, R., Alonso, J., Angermeyer, M., Demyttenaere, K., … Jin, R. (2007). Cross-national prevalence and correlates of adult attention-deficit hyperactivity disorder. *British Journal of Psychiatry*, *190*(MAY), 402–409. http://doi.org/10.1192/bjp.bp.106.034389

Frederick, J. (2012). Psychophysics of EEG alpha state discrimination. *Consciousness and*

Cognition, *21*(3), 1345–1354. http://doi.org/10.1016/j.concog.2012.06.009.PSYCHOPHYSICS

Fusar-Poli, P., Rubia, K., Rossi, G., Sartori, G., & Balottin, U. (2012). Striatal dopamine transporter alterations in ADHD: Pathophysiology or adaptation to psychostimulants? A meta-analysis. *American Journal of Psychiatry*. http://doi.org/10.1176/appi.ajp.2011.11060940

Gani, C. (2009). Long term effects after feedback of slow cortical potentials and of Theta / Beta - amplitudes in children with Attention Deficit Hyperactivity Disorder (ADHD) Long term effects after feedback of slow cortical potentials and of Theta / Beta - amplitudes.

Gevensleben, H., Albrecht, B., LÃ¼tcke, H., Auer, T., Dewiputri, W. I., Schweizer, R., … Rothenberger, A. (2014). Neurofeedback of slow cortical potentials: neural mechanisms and feasibility of a placebo-controlled design in healthy adults. *Frontiers in Human Neuroscience*, *8*(December), 1–13. http://doi.org/10.3389/fnhum.2014.00990

Gevensleben, H., Holl, B., Albrecht, B., Schlamp, D., Kratz, O., Studer, P., … Heinrich, H. (2009a). Distinct EEG effects related to neurofeedback training in children with ADHD: A randomized controlled trial. *International Journal of Psychophysiology*. http://doi.org/10.1016/j.ijpsycho.2009.08.005

Gevensleben, H., Holl, B., Albrecht, B., Schlamp, D., Kratz, O., Studer, P., … Heinrich, H. (2009b). Distinct EEG effects related to neurofeedback training in children with ADHD: a randomized controlled trial. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *74*(2), 149–57. http://doi.org/10.1016/j.ijpsycho.2009.08.005

Gevensleben, H., Holl, B., Albrecht, B., Schlamp, D., Kratz, O., Studer, P., … Heinrich, H. (2010). Neurofeedback training in children with ADHD: 6-month follow-up of a randomised controlled trial. *European Child & Adolescent Psychiatry*, *19*(9), 715–24. http://doi.org/10.1007/s00787-010-0109-5

Gevensleben, H., Kleemeyer, M., Rothenberger, L. G., Studer, P., Flaig-Röhr, A., Moll, G. H., … Heinrich, H. (2014a). Neurofeedback in ADHD: Further pieces of the puzzle. *Brain Topography*, *27*(1), 20–32. http://doi.org/10.1007/s10548-013-0285-y

Gevensleben, H., Kleemeyer, M., Rothenberger, L. G., Studer, P., Flaig-Röhr, A., Moll, G. H., … Heinrich, H. (2014b). Neurofeedback in ADHD: further pieces of the puzzle. *Brain Topography*, *27*(1), 20–32. http://doi.org/10.1007/s10548-013-0285-y

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Behavior rating inventory of executive function. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *6*(3), 235–8. http://doi.org/10.1076/chin.6.3.235.3152

Gioia, G. a, Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Behavior rating inventory of executive function. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *6*(3), 235–8. http://doi.org/10.1076/chin.6.3.235.3152

Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *J Child Psychol Psychiatry*, *41*(5), 645–655. http://doi.org/10.1111/j.1469-7610.2000.tb02345.x

Gruzelier, J. H. (2013). EEG-neurofeedback for optimising performance. II: Creativity, the performing arts and ecological validity. *Neuroscience and Biobehavioral Reviews*, *44*, 142–158. http://doi.org/10.1016/j.neubiorev.2013.11.004

Gruzelier, J. H. (2014). EEG-neurofeedback for optimising performance. III: A review of methodological and theoretical considerations. *Neuroscience and Biobehavioral Reviews*, *44*, 159–182. http://doi.org/10.1016/j.neubiorev.2014.03.015

Gruzelier, J., Inoue, A., Smart, R., Steed, A., & Steffert, T. (2010). Acting performance and flow state enhanced with sensory-motor rhythm neurofeedback comparing ecologically valid immersive VR and training screen scenarios. *Neuroscience Letters*, *480*(2), 112–6. http://doi.org/10.1016/j.neulet.2010.06.019

Guger, C., Edlinger, G., Harkam, W., Niedermayer, I., & Pfurtscheller, G. (2003). How many people are able to operate an EEG-based brain-computer interface (BCI)? *IEEE Transactions on Neural Systems and Rehabilitation Engineering : A Publication of the IEEE Engineering in Medicine and Biology Society*, *11*(2), 145–7. http://doi.org/10.1109/TNSRE.2003.814481

Halder, S., Varkuti, B., Bogdan, M., Kübler, a, Rosenstiel, W., Sitaram, R., & Birbaumer, N. (2013). Prediction of brain-computer interface aptitude from individual brain structure. *Frontiers in Human Neuroscience*, *7*(April), 105. http://doi.org/10.3389/fnhum.2013.00105

Hanslmayr, S., Sauseng, P., Doppelmayr, M., Schabus, M., & Klimesch, W. (2005). Increasing Individual Upper Alpha Power by Neurofeedback Improves Cognitive Performance in Human Subjects. *Applied Psychophysiology and Biofeedback*, *30*(1), 1–10. http://doi.org/10.1007/s10484-005-2169-8

Hardt, J. V, & Kamiya, J. (1978). Anxiety change through electroencephalographic alpha feedback seen only in high anxiety subjects. *Science (New York, N.Y.)*, *201*, 79–81. http://doi.org/10.1126/science.663641

Heinrich, H., Gevensleben, H., Freisleder, F. J., Moll, G. H., & Rothenberger, A. (2004a). Training of slow cortical potentials in attention-deficit/hyperactivity disorder: Evidence for positive behavioral and neurophysiological effects. *Biological Psychiatry*, *55*(7), 772–775. http://doi.org/10.1016/j.biopsych.2003.11.013

Heinrich, H., Gevensleben, H., Freisleder, F. J., Moll, G. H., & Rothenberger, A. (2004b). Training of slow cortical potentials in attention-deficit/hyperactivity disorder: evidence for positive behavioral and neurophysiological effects. *Biological Psychiatry*, *55*(7), 772–5. http://doi.org/10.1016/j.biopsych.2003.11.013

Heywood, C., & Beale, I. (2003). EEG biofeedback vs. placebo treatment for attention-deficit/hyperactivity disorder: a pilot study. *Journal of Attention Disorders*, *7*(1), 43–55. http://doi.org/10.1177/108705470300700105

Hillard, B., El-Baz, A. S., Sears, L., Tasman, A., & Sokhadze, E. M. (2013a). Neurofeedback training aimed to improve focused attention and alertness in children with ADHD: a study of relative power of EEG rhythms using custom-made software application. *Clinical EEG and Neuroscience*, *44*(3), 193–202. http://doi.org/10.1177/1550059412458262

Hillard, B., El-Baz, A. S., Sears, L., Tasman, A., & Sokhadze, E. M. (2013b). Neurofeedback Training Aimed to Improve Focused Attention and Alertness in Children With ADHD A Study of Relative Power of EEG Rhythms Using Custom-Made Software Application. *Clinical EEG and Neuroscience*, *44*(3), 193–202. http://doi.org/10.1177/1550059412458262

Hinterberger, T., Klinger, C., Strehl, U., Birbaumer, N., Goth, G., & Leins, U. (2006). Self-regulation of Slow Cortical Potentials: A New Treatment for Children With Attention-Deficit/Hyperactivity Disorder. *Pediatrics*, *118*(5), e1530-1540. http://doi.org/10.1542/peds.2005-2478

Hodgson, K., Hutchinson, A. D., & Denson, L. (2014). Nonpharmacological treatments for

ADHD: a meta-analytic review. *Journal of Attention Disorders*, *18*(4), 275–82. http://doi.org/10.1177/1087054712444732

Hoedlmoser, K., Pecherstorfer, T., Gruber, G., Anderer, P., Doppelmayr, M., Klimesch, W., & Schabus, M. (2008). Instrumental conditioning of human sensorimotor rhythm (12-15 Hz) and its impact on sleep as well as declarative learning. *Sleep*, *31*(10), 1401–8. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2572745&tool=pmcentrez&rendertype=abstract

Holtmann, M., Grasmann, D., Cionek-Szpak, E., Hager, V., Panzner, N., Beyer, A., … Stadler, C. (2009). Spezifische Wirksamkeit von Neurofeedback auf die Impulsivität bei ADHS. *Kindheit Und Entwicklung*, *18*(2), 95–104. http://doi.org/10.1026/0942-5403.18.2.95

Holtmann, M., Sonuga-Barke, E., Cortese, S., & Brandeis, D. (2014). Neurofeedback for Attention-Deficit/Hyperactivity Disorder. *Child and Adolescent Psychiatric Clinics of North America*. http://doi.org/10.1016/j.chc.2014.05.006

Huang, Y.-S., Tsai, M.-H., & Guilleminault, C. (2011). Pharmacological Treatment of ADHD and the Short and Long Term Effects on Sleep. *Current Pharmaceutical Design*, *17*, 1450–1458. http://doi.org/10.2174/138161211796197179

Hurt, E., Arnold, L. E., & Lofthouse, N. (2014). Quantitative EEG neurofeedback for the treatment of pediatric attention-deficit/hyperactivity disorder, autism spectrum disorders, learning disorders, and epilepsy. *Child and Adolescent Psychiatric Clinics of North America*. http://doi.org/10.1016/j.chc.2014.02.001

Janssen, T. W. P., Bink, M., Geladé, K., Van Mourik, R., Maras, A., & Oosterlaan, J. (2016). A randomized controlled trial into the effects of neurofeedback, methylphenidate, and physical activity on EEG power spectra in children with ADHD. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *57*(5), 633–644. http://doi.org/10.1111/jcpp.12517

Janssen, T. W. P., Bink, M., Weeda, W. D., Gelad??, K., van Mourik, R., Maras, A., & Oosterlaan, J. (2016). Learning curves of theta/beta neurofeedback in children with ADHD. *European Child and Adolescent Psychiatry*, 1–10. http://doi.org/10.1007/s00787-016-0920-8

Jensen, P. S., Arnold, L. E., Swanson, J. M., Vitiello, B., Abikoff, H. B., Greenhill, L. L., … Hur, K. (2007). 3-year follow-up of the NIMH MTA study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*(8), 989–1002. http://doi.org/10.1097/CHI.0b013e3180686d48

Keizer, A. W., Verment, R. S., & Hommel, B. (2010). Enhancing cognitive control through neurofeedback: a role of gamma-band activity in managing episodic retrieval. *NeuroImage*, *49*(4), 3404–13. http://doi.org/10.1016/j.neuroimage.2009.11.023

Kerson, C. (2013). A proposed multisite double-blind randomized clinical trial of neurofeedback for ADHD: need, rationale, and strategy. *Journal of Attention Disorders*, *17*(5), 420–36. http://doi.org/10.1177/1087054713482580

Kessler, R., Adler, L. A., & al., et. (2006). The prevalence and correlates of adult ADHD in the United States: results from the National Comorbidity Survey Replication. *American Journal of …*. Retrieved from http://ajp.psychiatryonline.org/cgi/content/abstract/ajp%5Cn163/4/716%5Cnfile://localhost/Users/Bonnie/Dropbox/Papers/2006/Kessler/American Journal of … 2006 Kessler.pdf%5Cnpapers://06eccb9f-3e0f-4128-8f71-2abf6c034945/Paper/p1598

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, *29*(2–3), 169–195.

http://doi.org/10.1016/S0165-0173(98)00056-3

Klimesch, W., Doppelmayr, M., & Hanslmayr, S. (2006). Chapter 10 Upper alpha ERD and absolute power: their meaning for memory performance. *Progress in Brain Research*, *159*, 151–165. http://doi.org/10.1016/S0079-6123(06)59010-7

Kober, S. E., Witte, M., Ninaus, M., Neuper, C., & Wood, G. (2013). Learning to modulate one's own brain activity: the effect of spontaneous mental strategies. *Frontiers in Human Neuroscience*, *7*(October), 695. http://doi.org/10.3389/fnhum.2013.00695

Kotchoubey, B., Blankenhorn, V., Fröscher, W., Strehl, U., & Birbaumer, N. (1997). Stability of cortical self-regulation in epilepsy patients. *Neuroreport*, *8*(8), 1867–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9223067

Kotchoubey, B., Kübler, A., Strehl, U., Flor, H., & Birbaumer, N. (2002). Can humans perceive their brain states? *Consciousness and Cognition*, *11*(1), 98–113. http://doi.org/10.1006/ccog.2001.0535

Kotchoubey, B., Strehl, U., Holzapfel, S., Blankenhorn, V., Fröscher, W., & Birbaumer, N. (1999). Negative potential shifts and the prediction of the outcome of neurofeedback therapy in epilepsy. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *110*(4), 683–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10378738

Kouijzer, M. E. J., van Schie, H. T., Gerrits, B. J. L., Buitelaar, J. K., & de Moor, J. M. H. (2013). Is EEG-biofeedback an effective treatment in autism spectrum disorders? A randomized controlled trial. *Applied Psychophysiology and Biofeedback*, *38*(1), 17–28. http://doi.org/10.1007/s10484-012-9204-3

Kropotov, J. D., Grin-Yatsenko, V., Ponomarev, V., Chutko, L., Yakovenko, E., & Nikishena, I. (2005). ERPs correlates of EEG relative beta training in ADHD children. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *55*(1), 23–34. http://doi.org/10.1016/j.ijpsycho.2004.05.011

Kübler, A., Mattia, D., George, H., Doron, B., & Neuper, C. (2007). How much learning is involved in BCI-control? *Biol*, *109*(2002), 224631. Retrieved from http://www.tobi-project.org/sites/default/files/public/Publications/How_much_learning_is_involved_in_BCI-control.pdf

Kübler, A., & Neumann, N. (2004). Predictability of Brain-Computer Communication. *Journal of …*, *18*, 121–129. http://doi.org/10.1027/0269-8803.18.2

Lansbergen, M. M., van Dongen-Boomsma, M., Buitelaar, J. K., & Slaats-Willemse, D. (2011). ADHD and EEG-neurofeedback: a double-blind randomized placebo-controlled feasibility study. *Journal of Neural Transmission (Vienna, Austria : 1996)*, *118*(2), 275–84. http://doi.org/10.1007/s00702-010-0524-2

Leins, U., Goth, G., Hinterberger, T., Klinger, C., Rumpf, N., & Strehl, U. (2007a). Neurofeedback for children with ADHD: A comparison of SCP and Theta/Beta protocols. *Applied Psychophysiology Biofeedback*, *32*(2), 73–88. http://doi.org/10.1007/s10484-007-9031-0

Leins, U., Goth, G., Hinterberger, T., Klinger, C., Rumpf, N., & Strehl, U. (2007b). Neurofeedback for children with ADHD: a comparison of SCP and Theta/Beta protocols. *Applied Psychophysiology and Biofeedback*, *32*(2), 73–88. http://doi.org/10.1007/s10484-007-9031-0

Lidzba, K. , Chritiansen, H. , Drechsler, R. (2013). *Conners Skalen zu Aufmerksamkeit und Verhalten-3. Deutschsprachige Adaptation der Conners-3rd edition von Keith Conners .* (Hube). Bern.

Liechti, M. D., Maurizio, S., Heinrich, H., Jäncke, L., Meier, L., Steinhausen, H.-C., … Brandeis, D. (2012). First clinical trial of tomographic neurofeedback in attention-deficit/hyperactivity disorder: evaluation of voluntary cortical control. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *123*(10), 1989–2005. http://doi.org/10.1016/j.clinph.2012.03.016

Liechti, M. D., Valko, L., Müller, U. C., Döhnert, M., Drechsler, R., Steinhausen, H. C., & Brandeis, D. (2013). Diagnostic value of resting electroencephalogram in attention-deficit/ hyperactivity disorder across the lifespan. *Brain Topography*, *26*, 135–151. http://doi.org/10.1007/s10548-012-0258-6

Liotti, M., Pliszka, S. R., Perez, R., Luus, B., Glahn, D., & Semrud-Clikeman, M. (2007). Electrophysiological correlates of response inhibition in children and adolescents with ADHD: Influence of gender, age, and previous treatment history. *Psychophysiology*, *44*(6), 936–948. http://doi.org/10.1111/j.1469-8986.2007.00568.x

Lofthouse, N., Arnold, L. E., Hersch, S., Hurt, E., & DeBeus, R. (2012). A review of neurofeedback treatment for pediatric ADHD. *Journal of Attention Disorders*, *16*(5), 351–72. http://doi.org/10.1177/1087054711427530

Logemann, H. N. A., Lansbergen, M. M., Van Os, T. W. D. P., Böcker, K. B. E., & Kenemans, J. L. (2010). The effectiveness of EEG-feedback on attention, impulsivity and EEG: a sham feedback controlled study. *Neuroscience Letters*, *479*(1), 49–53. http://doi.org/10.1016/j.neulet.2010.05.026

Loo, S. K., Teale, P. D., & Reite, M. L. (1999). EEG correlates of methylphenidate response among children with ADHD: a preliminary report - A window in prefrontal function in children. *Biological Psychiatry*, *45*(12), 1657–1660. http://doi.org/10.1016/S0006-3223(98)00250-9

Loo, S., & Makeig, S. (2012). Clinical Utility of EEG in Attention-Deficit/Hyperactivity Disorder: A Research Update. *Neurotherapeutics*, *9*(3), 569–587. http://doi.org/10.1007/s13311-012-0131-z

Lotte, F., Larrue, F., & Mühl, C. (2013). Flaws in current human training protocols for spontaneous Brain-Computer Interfaces: lessons learned from instructional design. *Frontiers in Human Neuroscience*, *7*(September), 568. http://doi.org/10.3389/fnhum.2013.00568

Lubar. (1995). Evaluation of the Effectiveness of EEG.., 1–24. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:No+Title#0

Lubar, J. F., Swartwood, M. O., Swartwood, J. N., & O'Donnell, P. H. (1995). Evaluation of the effectiveness of EEG neurofeedback training for ADHD in a clinical setting as measured by changes in T.O.V.A. scores, behavioral ratings, and WISC-R performance. *Biofeedback and Self-Regulation*, *20*, 83–99. http://doi.org/10.1007/BF01712768

Lutzenberger, W., Birbaumer, N., Elbert, T., & Rockstroh, B. (1982). BIOFEEDBACK PRODUCED SLOW BRAIN POTENTIALS, *14*, 99–111.

Matousek, M., & Petersén, I. (1983). A method for assessing alertness fluctuations from EEG spectra. *Electroencephalography and Clinical Neurophysiology*, *55*(1), 108–113. http://doi.org/10.1016/0013-4694(83)90154-2

Matthis, P., Scheffner, D., & Benninger, C. (1981). Spectral analysis of the EEG: Comparison of various spectral parameters. *Electroencephalography and Clinical Neurophysiology*, *52*(2), 218–221. http://doi.org/10.1016/0013-4694(81)90171-1

Maurizio, S., Liechti, M. D., Heinrich, H., Jäncke, L., Steinhausen, H.-C., Walitza, S., … Drechsler, R. (2014a). Comparing tomographic EEG neurofeedback and EMG

biofeedback in children with attention-deficit/hyperactivity disorder. *Biological Psychology*, *95*, 31–44. http://doi.org/10.1016/j.biopsycho.2013.10.008

Maurizio, S., Liechti, M. D., Heinrich, H., Jäncke, L., Steinhausen, H.-C., Walitza, S., … Drechsler, R. (2014b). Comparing tomographic EEG neurofeedback and EMG biofeedback in children with attention-deficit/hyperactivity disorder. *Biological Psychology*, *95*, 31–44. http://doi.org/10.1016/j.biopsycho.2013.10.008

Mayer, K., Wyckoff, S. N., Schulz, U., & Strehl, U. (2012). Neurofeedback for Adult Attention-Deficit/Hyperactivity Disorder: Investigation of Slow Cortical Potential Neurofeedback—Preliminary Results. *Journal of Neurotherapy*, *16*(1), 37–45. http://doi.org/10.1080/10874208.2012.650113

Mazzone, L., Postorino, V., Reale, L., Guarnera, M., Mannino, V., Armando, M., … Vicari, S. (2013). Self-Esteem Evaluation in Children and Adolescents Suffering from ADHD. *Journal List Clin Pract Epidemiol Ment Health v. Clin Pract Epidemiol Ment Health*, *9*(9), 96–102. http://doi.org/10.2174/1745017901309010096

Monastra, V. J., Monastra, D. M., & George, S. (2002). The effects of stimulant therapy, EEG biofeedback, and parenting style on the primary symptoms of attention-deficit/hyperactivity disorder. *Applied Psychophysiology and Biofeedback*, *27*(4), 231–49. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12557451

Monto, S., Palva, S., Voipio, J., & Palva, J. M. (2008). Very slow EEG fluctuations predict the dynamics of stimulus detection and oscillation amplitudes in humans. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(33), 8268–8272. http://doi.org/10.1523/JNEUROSCI.1910-08.2008

Moore, A., Gruber, T., Derose, J., & Malinowski, P. (2012). Regular, brief mindfulness meditation practice improves electrophysiological markers of attentional control. *Frontiers in Human Neuroscience*, *6*, 18. http://doi.org/10.3389/fnhum.2012.00018

Nan, W., Rodrigues, J. P., Ma, J., Qu, X., Wan, F., Mak, P.-I., … Rosa, A. (2012). Individual alpha neurofeedback training effect on short term memory. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *86*(1), 83–7. http://doi.org/10.1016/j.ijpsycho.2012.07.182

Nan, W., Wan, F., Vai, M. I., & Da Rosa, A. C. (2015). Resting and Initial Beta Amplitudes Predict Learning Ability in Beta/Theta Ratio Neurofeedback Training in Healthy Young Adults. *Frontiers in Human Neuroscience*, *9*(December), 677. http://doi.org/10.3389/fnhum.2015.00677

Neumann, N., & Birbaumer, N. (2003). Predictors of successful self control during brain-computer communication. *Journal of Neurology, Neurosurgery, and Psychiatry*, *74*(8), 1117–21. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1738580&tool=pmcentrez&rendertype=abstract

Newark, P. E., & Stieglitz, R. D. (2010). Therapy-relevant factors in adult ADHD from a cognitive behavioural perspective. *ADHD Attention Deficit and Hyperactivity Disorders*. http://doi.org/10.1007/s12402-010-0023-1

Nijboer, F., Furdea, A., Gunst, I., Mellinger, J., McFarland, D. J., Birbaumer, N., & Kübler, A. (2008). An auditory brain-computer interface (BCI). *Journal of Neuroscience Methods*, *167*, 43–50. http://doi.org/10.1016/j.jneumeth.2007.02.009

Oerter, R. & Montada, L. (Hrsg.). (1995). *Entwicklungspsychologie. Ein Lehrbuch*. (L. Oerter, R. & Montada, Ed.). Weinheim.

Owens, J. S., Goldfine, M. E., Evangelista, N. M., Hoza, B., & Kaiser, N. M. (2007). A critical

review of self-perceptions and the positive illusory bias in children with ADHD. *Clinical Child and Family Psychology Review*, *10*(4), 335–351. http://doi.org/10.1007/s10567-007-0027-3

Pfurtscheller, G., Stancák, A., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band - An electrophysiological correlate of cortical idling: A review. *International Journal of Psychophysiology*, *24*, 39–46. http://doi.org/10.1016/S0167-8760(96)00066-9

Pineda, J. A., Carrasco, K., Datko, M., Pillen, S., & Schalles, M. (2014). Neurofeedback training produces normalization in behavioural and electrophysiological measures of high-functioning autism. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*, 20130183. http://doi.org/10.1098/rstb.2013.0183

Plotkin, W. P., & Rice, K. M. (1981). *Biofeedback as a placebo: anxiety reduction facilitated by training in either suppression or enhancement of alpha brainwaves. Journal of consulting and clinical psychology* (Vol. 49).

Polanczyk, G., De Lima, M. S., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: A systematic review and metaregression analysis. *American Journal of Psychiatry*, *164*(6), 942–948. http://doi.org/10.1176/appi.ajp.164.6.942

Pop-Jordanova, N., Markovska-Simoska, S., & Zorcec, T. (2005). Neurofeedback treatment of children with attention deficit hyperactivity disorder. *Prilozi / Makedonska Akademija Na Naukite I Umetnostite, Oddelenie Za Biološki I Medicinski Nauki = Contributions / Macedonian Academy of Sciences and Arts, Section of Biological and Medical Sciences*, *26*(1), 71–80. http://doi.org/16118616

Rockstroh, B., Elbert, T., Birbaumer, N., & Lutzenberger, W. (1990). Biofeedback-produced hemispheric asymmetry of slow cortical potentials and its behavioural effects. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *9*(2), 151–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2228749

Ros, T., Michela, A., Bellman, A., Vuadens, P., & Saj, A. (2017). Increased Alpha-Rhythm Dynamic Range Promotes Recovery from Visuospatial Neglect : A Neurofeedback Study, *2017*(January). http://doi.org/10.1155/2017/7407241

Ros, T., Moseley, M. J., Bloom, P. a, Benjamin, L., Parkinson, L. a, & Gruzelier, J. H. (2009). Optimizing microsurgical skills with EEG neurofeedback. *BMC Neuroscience*, *10*, 87. http://doi.org/10.1186/1471-2202-10-87

Russell-Chapin, L., Kemmerly, T., Liu, W.-C., Zagardo, M. T., Chapin, T., Dailey, D., & Dinh, D. (2013). The Effects of Neurofeedback in the Default Mode Network: Pilot Study Results of Medicated Children with ADHD. *Journal of Neurotherapy*, *17*(1), 35–42. http://doi.org/10.1080/10874208.2013.759017

Schabus, M., Heib, D. P. J., Lechinger, J., Griessenberger, H., Klimesch, W., Pawlizki, A., … Hoedlmoser, K. (2014). Enhancing sleep quality and memory in insomnia using instrumental sensorimotor rhythm conditioning. *Biological Psychology*, *95*, 126–34. http://doi.org/10.1016/j.biopsycho.2013.02.020

Schoenberg, P. L. A., Hepark, S., Kan, C. C., Barendregt, H. P., Buitelaar, J. K., & Speckens, A. E. M. (2014). Effects of mindfulness-based cognitive therapy on neurophysiological correlates of performance monitoring in adult attention-deficit/hyperactivity disorder. *Clinical Neurophysiology*, *125*, 1407–1416. http://doi.org/10.1016/j.clinph.2013.11.031

Sonuga-Barke, E., Brandeis, D., Holtmann, M., & Cortese, S. (2014). Computer-based Cognitive Training for ADHD. A Review of Current Evidence. *Child and Adolescent Psychiatric Clinics of North America*, *23*(4), 807–824.

http://doi.org/10.1016/j.chc.2014.05.009

Sonuga-Barke, Brandeis, D., Ph, D., Cortese, S., Ph, D., Daley, D., … Ph, D. (2013). Nonpharmacological Interventions for ADHD: Systematic Review and Meta-Analyses of Randomized Controlled Trials of Dietary and Psychological Treatments, (March), 275–289.

Spruyt, K., & Gozal, D. (2011). Sleep disturbances in children with attention-deficit/hyperactivity disorder. *Expert Review of Neurotherapeutics*, *11*(4), 565–77. http://doi.org/10.1586/ern.11.7

Steiner, N. J., Frenette, E. C., Rene, K. M., Brennan, R. T., & Perrin, E. C. (2014). In-school neurofeedback training for ADHD: sustained improvements from a randomized control trial. *Pediatrics*, *133*(3), 483–92. http://doi.org/10.1542/peds.2013-2059

Sterman, M. B., Howe, R. C., & Macdonald, L. R. (1970). Facilitation of spindle-burst sleep by conditioning of electroencephalographic activity while awake. *Science (New York, N.Y.)*, *167*, 1146–1148. http://doi.org/10.1126/science.167.3921.1146

Strehl, U., Kotchoubey, B., Trevorrow, T., & Birbaumer, N. (2005). Predictors of seizure reduction after self-regulation of slow cortical potentials as a treatment of drug-resistant epilepsy. *Epilepsy & Behavior : E&B*, *6*(2), 156–66. http://doi.org/10.1016/j.yebeh.2004.11.004

Strehl, U., Leins, U., Goth, G., Klinger, C., Hinterberger, T., & Birbaumer, N. (2006a). Self-regulation of Slow Cortical Potentials: A New Treatment for Children With Attention-Deficit/Hyperactivity Disorder. *Pediatrics*, *118*(5), e1530–e1540. http://doi.org/10.1542/peds.2005-2478

Strehl, U., Leins, U., Goth, G., Klinger, C., Hinterberger, T., & Birbaumer, N. (2006b). Self-regulation of slow cortical potentials: a new treatment for children with attention-deficit/hyperactivity disorder. *Pediatrics*, *118*(5), e1530-40. http://doi.org/10.1542/peds.2005-2478

Swartwood, J. N., Swartwood, M. O., Lubar, J. F., & Timmermann, D. L. (2003). EEG differences in ADHD-combined type during baseline and cognitive tasks. *Pediatric Neurology*, *28*(3), 199–204. http://doi.org/10.1016/S0887-8994(02)00514-3

Swartwood, M. O., Swartwood, J. N., Lubar, J. F., Timmermann, D. L., Zimmerman, A. W., & Muenchen, R. A. (1998). Methylphenidate effects on EEG, behavior, and performance in boys with ADHD. *Pediatric Neurology*, *18*(3), 244–250. http://doi.org/10.1016/S0887-8994(97)00205-1

Thompson, L., & Thompson, M. (1998). Neurofeedback combined with training in metacognitive strategies: effectiveness in students with ADD. *Applied Psychophysiology and Biofeedback*, *23*(4), 243–63. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10457815

Trautner, H. M. (1991). *Lehrbuch der Entwicklungspsychologie*. *Grundlagen und Methoden*.

Van Dongen-Boomsma, M., Vollebregt, M. A., Slaats-Willemse, D., & Buitelaar, J. K. (2013). A randomized placebo-controlled trial of electroencephalographic (EEG) neurofeedback in children with attention-deficit/hyperactivity disorder. *Journal of Clinical Psychiatry*, *74*(8), 821–827. http://doi.org/10.4088/JCP.12m08321

Van Doren, J., Heinrich, H., Bezold, M., Reuter, N., Kratz, O., Horndasch, S., … Studer, P. (2016). Theta/beta neurofeedback in children with ADHD: Feasibility of a short-term setting and plasticity effects. *International Journal of Psychophysiology*, *112*(November), 80–88. http://doi.org/10.1016/j.ijpsycho.2016.11.004

Vidaurre, C., & Blankertz, B. (2010). Towards a cure for BCI illiteracy. *Brain Topography*, *23*(2),

194–8. http://doi.org/10.1007/s10548-009-0121-6

Vollebregt, M. a, van Dongen-Boomsma, M., Buitelaar, J. K., & Slaats-Willemse, D. (2014). Does EEG-neurofeedback improve neurocognitive functioning in children with attention-deficit/hyperactivity disorder? A systematic review and a double-blind placebo-controlled study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *55*(5), 460–72. http://doi.org/10.1111/jcpp.12143

Waldmann, H. C. (2008). Kurzformen des HAWIK-IV: Statistische bewertung in verschiedenen anwendungsszenarien. *Diagnostica*, *54*(4), 202–210. http://doi.org/10.1026/0012-1924.54.4.202

Wan, F., Nan, W., Vai, M., & Rosa, A. (2014). Resting alpha activity predicts learning ability in alpha neurofeedback. *Name: Frontiers in Human …*. http://doi.org/10.3389/fnhum.2014.00500

Wangler, S., Gevensleben, H., Albrecht, B., Studer, P., Rothenberger, A., Moll, G. H., & Heinrich, H. (2011). Neurofeedback in children with ADHD: Specific event-related potential findings of a randomized controlled trial. *Clinical Neurophysiology*, *122*(5), 942–950. http://doi.org/10.1016/j.clinph.2010.06.036

Weber, E., Köberl, a, Frank, S., & Doppelmayr, M. (2011). Predicting successful learning of SMR neurofeedback in healthy participants: methodological considerations. *Applied Psychophysiology and Biofeedback*, *36*(1), 37–45. http://doi.org/10.1007/s10484-010-9142-x

Willcutt, E. G., Petrill, S. A., Wu, S., Boada, R., Defries, J. C., Olson, R. K., & Pennington, B. F. (2013). Comorbidity between reading disability and math disability: Concurrent psychopathology, functional impairment, and neuropsychological functioning. *Journal of Learning Disabilities*, *46*(6), 500–16. http://doi.org/10.1177/0022219413477476

Witte, M., Kober, S. E., Ninaus, M., Neuper, C., & Wood, G. (2013). Control beliefs can predict the ability to up-regulate sensorimotor rhythm during neurofeedback training. *Frontiers in Human Neuroscience*, *7*(August), 478. http://doi.org/10.3389/fnhum.2013.00478

Zoefel, B., Huster, R. J., & Herrmann, C. S. (2011). Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance. *NeuroImage*, *54*(2), 1427–31. http://doi.org/10.1016/j.neuroimage.2010.08.078

Zuberer, A., Brandeis, D., & Drechsler, R. (2015). Are treatment effects of neurofeedback training in children with ADHD related to the successful regulation of brain activity ? A review on the learning of regulation of brain activity and a contribution to the discussion Are treatment effects of neurofeed, *9*(March), 1–15. http://doi.org/10.3389/fnhum.2015.00135

8 Appendices

8.1 Appendix A

Table S1. Results for linear mixed effects models for cross-session NF learning with task type (Feedback / Transfer) in titles.

	Feedback activation			Feedback deactivation			Transfer activation			Transfer deactivation		
	B	*CI*	*p*	*B*	*CI*	*p*	*B*	*CI*	*p*	*B*	*CI*	*p*
Fixed Parts												
Intercept	-3.53	-5.27 – -1.79	**<.001**	-2.47	-3.65 – -1.28	**<.001**	-1.26	-2.89 – 0.36	.134	-4.40	-5.95 – -2.86	**<.001**
Session	-0.02	-0.19 – 0.14	.792	0.19	0.03 – 0.35	**.021**	0.03	-0.15 – 0.21	.753	0.30	0.12 – 0.47	**.002**
IQ	-0.10	-0.19 – -0.01	**.031**	-0.07	-0.12 – -0.01	**.034**	-0.07	-0.13 – 0.00	.058			
Age	-0.87	-1.49 – -0.24	**.009**	-1.01	-1.41 – -0.60	**<.001**	-0.76	-1.22 – -0.29	**.003**	-0.61	-1.02 – -0.19	**.006**
MPH							4.55	1.74 – 7.36	**.003**			
Session:MPH							-0.27	-0.58 – 0.04	.094			
MPHd										0.60	0.11 – 1.09	**.019**
Random Parts												
σ^2		35.818			34.203			49.595			56.338	
$\tau_{00,\ subject}$		26.933			7.217			6.952			10.956	
ρ_{01}		-0.608			-0.585			-0.347			-0.769	
R^2 / Ω_0^2		.462 / .452			.355 / .341			.271 / .256			.239 / .214	

Feedback/Transfer: Condition where a feedback stimulus is (feedback) or is not (transfer) involved. *Task:* Performance in the deactivation (generation of positive potential shifts) versus activation Task (Generation of negative potential shifts). *MPH*: On regular metylphenidate medication (yes versus no). *MPHd*: intake duration in years. σ^2: Within-subject residual variance. *t00, subject*: Between-subject variance. τ_{11}: Random slope variance.ρ_{01}: Random-Intercept-Slope correlation.

8.2 Appendix B

S1. Description of Participants.

N (total)	44
Male/female (N)	25/19
Setting school/clinic (N)	21/23
Age (years)	11.3 ± 2.1
Methylphenidate (with/without)	14/20
Dose (mg)	23.0 ± 12.1
Duration intake (years)	2.6 ± 2.6
IQ	108.8 ± 15.3

S2. Effects considered for statistical analysis.

Model specifications	***Measure***
Condition	
Feedback (FB)	Online feedback stimulus visible.
Transfer (TR)	No online feedback.
Effects	
session	Training was applied across 15 training double sessions in total.
Task	Deactivation (generation of positive potential shifts of SCPs) vs. Activation (generation of negative potential shifts of SCPs).
Inter-session- interval	Days passed between training sessions.
Age	In years (continuous variable in the model. Only for visualization in plots dichotomized into younger and older age classes).
MPH	Being on regular stimulant medication (methylphenidate), factorized into yes versus no.
Stimulants intake duration	Years of MPH intake.
Dosage of stimulant medication	Methylphenidate (MPH) in mg.
Sex	Factorized into female versus male.
IQ	Estimated IQ (WISC-IV short form).
Setting	Factorized into school setting versus clinical setting.
Severity of ADHD symptoms	T-values of the Conners 3 DSM-IV indices for Hyperactivity and Inattention based on parent and teacher ratings at screening.
Artifact rate	Percentage of rejected trials within a session.

S3. Results of cross-session NF learning

	Feedback			Transfer		
	B	*CI*	*p*	*B*	*CI*	*p*
Fixed Parts						
Intercept	1.48	-1.01 – 3.97	.249	-2.58	-4.76 – -0.41	**.024**
Session	0.17	-0.08 – 0.42	.194	0.24	0.02 – 0.46	**.041**
MPH	-1.87	-6.18 – 2.44	.398	-5.06	-8.82 – -1.31	**.011**
Age	-0.29	-1.53 – 0.95	.654			
Session : MPH	0.15	-0.29 – 0.58	.510	0.42	0.04 – 0.79	**.034**
Session : Age	-0.03	-0.16 – 0.09	.635			
MPH : Age	-1.86	-3.82 – 0.09	.067			
Session : MPH : Age	0.36	0.17 – 0.56	**.001**			
Random Parts						
σ^2		37.072			90.221	
$\tau_{00,\ subject}$		40.317			12.021	
$\tau_{01,\ Session}$		0.6146			0.2205	
ρ_{01}		-0.669			-0.700	
$N_{subject}$		48			48	
Observations		693			692	
R^2 / Ω_0^2		.533 / .525			.168 / .149	

Feedback/Transfer: Condition where a feedback stimulus is (feedback) or is not (transfer) involved. *Task:* Performance in the deactivation (generation of positive potential shifts) versus activation task (generation of negative potential shifts). *MPH*: On regular metyl-phenidate medication (yes versus no). σ^2: Within-subject residual variance. *t00, subject*: Between-subject-variance. τ_{11}: Random slope variance. ρ_{01}: Random intercept slope correlation.

S4. Effect plot for cross-session NF learning.

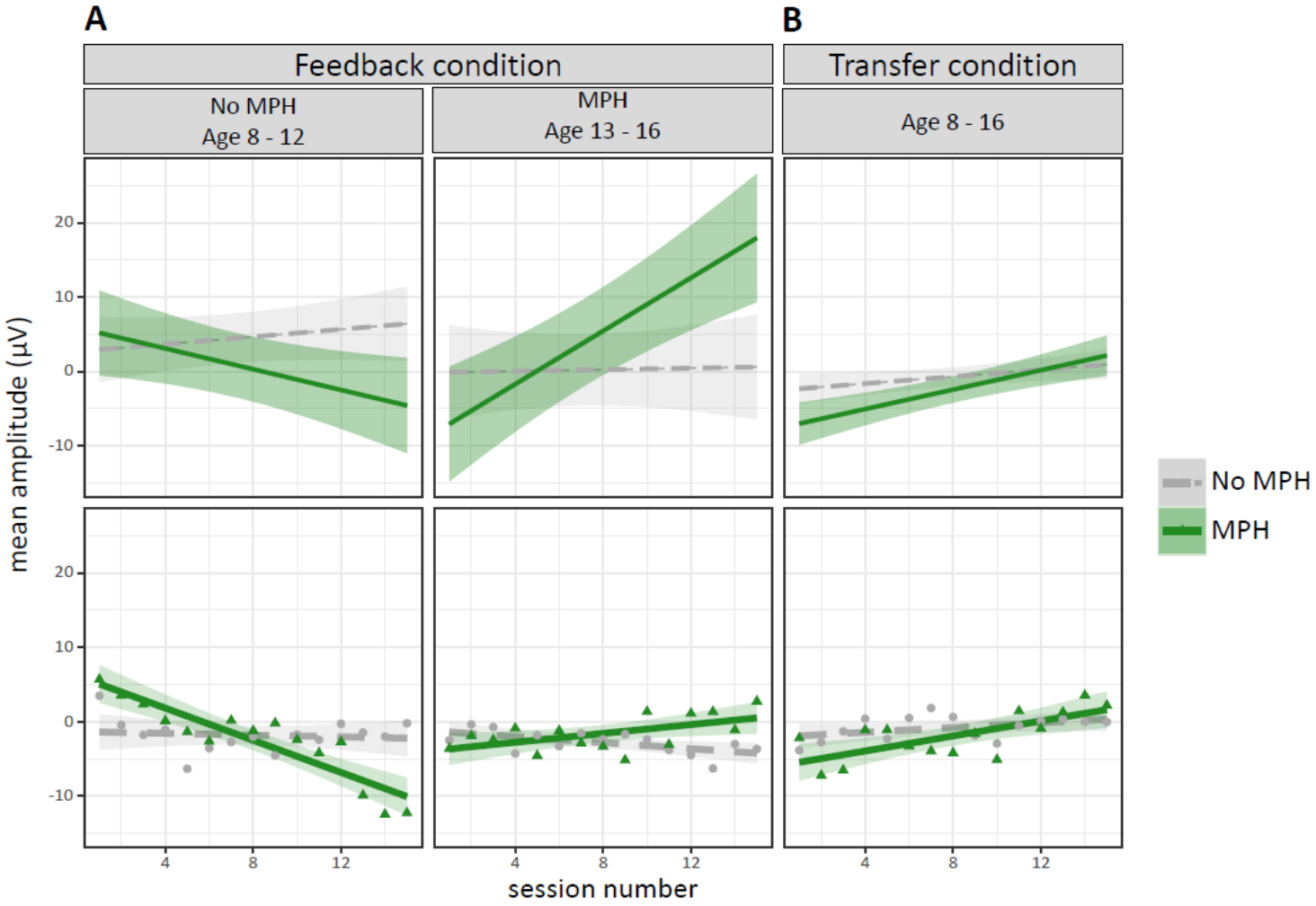

Figure S4. Visualization of cross- session NF-learningin the Feedback (A) and Transfer task (B). The dependent variable is the difference between mean amplitude (µV) of baseline corrected trials od both tasks (deactivation minus activation). For comparison between effects and raw data, see scatter plot under each effects panel, fitted with a fixed linear regression based on the same factors as in A. **A**: Interaction plot for the fixed effects session number, MPH and age. **B**: Interaction for the fixed effects session number and MPH. *Session number:* 15 sessions in total. *Condition:* Deactivation: Generation of positive potential shifts. Activation: Generation of negative potential shifts. *MPH:* being on regular methylphenidate medication (yes versus no). *Task:* Feedback: Feedback stimulus visible. Transfer: No feedback stimulus visible. **A and B**: For visualisation age is subdivided into two age classes (8-12 and 13-16 years), but preserved as a continuous variable in the original model.

S5. Results for linear mixed effects models for upper alpha power across sessions.

	Upper alpha power spectral density (Hz/mV^2)		
	B	*CI*	*p*
Fixed Parts			
Intercept	1.17	0.99 – 1.35	**<.001**
Session	0.01	0.01 – 0.02	**<.001**
BL2 vs. BL1	-0.11	-0.14 – -0.08	**<.001**
Age	-0.14	-0.22 – -0.06	**.002**
Okzipital vs. frontal	0.31	0.28 – 0.35	**<.001**
Random Parts			
σ^2		0.156	
$\tau_{00,\ subject}$		0.366	
ρ_{01}		0.361	

Session: session number (15 double sessions in total). *BL2 vs. BL1*: factor indicating time of recording, namely either before a session (BL1), or after a session (BL2).*Occipital vs. frontal*: Mean of electrodes Oz versus the mean of electrodes Cz and Fz. σ^2: Within-subject residual variance. ρ_{01}: Random intercept-slope correlation.*t00, subject*: between-subject variance.

Agnieszka Zuberer

Curriculum Vitae

Department of Child and Adolescent Psychiatry
University of Zürich
Eisengasse 16, 8008 Zürich, Switzerland
+41 779212940
azuberer[at]gmail[dot]com

Employment

Current Position

2017-Present **Adjunct Researcher**, *Department of Child and Adolescent Psychiatry*, Psychiatric University Hospital, University of Zürich, Zürich, Switzerland.

Previous Positions

2013-2017 **Pre-Doctoral Research Fellow**, *Department of Child and Adolescent Psychiatry*, Psychiatric University Hospital, University of Zürich, Zürich, Switzerland.

Education

2017 **Dr. des. Psychology**, *University of Zürich*, Zürich, Switzerland, *magna cum laude*.
Thesis title: Specificity of neurofeedback training in attention-deficit/hyperactivity disorder (ADHD) - Multilevel modeling of EEG-learning.

2013 **M.Sc. Psychology**, *University of Zürich*, Zürich, Switzerland.
Specialization in neuropsychology, minor in philosophy and pedagogy.

Presentations

Conference Presentations

Apr. 2017 **Zuberer, A.**, Drechsler, R., Minder, F., Brandeis, D., *Multilevel modeling of neurofeedback EEG-learning in children and adolescents with ADHD*, 6th world congress on ADHD, Vancouver, Canada. .

Oct. 2016 **Zuberer, A.** Brandeis, D., Drechsler, R., *Knowing and changing mental states: What can be learned from performance on a self-appraisal task*, EUNTEHYDIS international conference on ADHD, Berlin, Germany.

Mar. 2016 **Zuberer, A.**Brandeis, D., Drechsler, R., *Self-regulated brain activity during neurofeedback training in children with ADHD*, Mind, Brain and Body Symposium, Berlin, Germany.

Oct. 2015 **Zuberer, A.** Brandeis, D., Drechsler, R., *Neurofeedback – different measures for training performance*, Burghölzli Psychiatry meeting, Zürich, Switzerland..

Oct. 2015 **Zuberer, A.** Brandeis, D., Drechsler, R., *Learning to regulate slow cortical potentials: A study on the diversity of learning patterns and possible mediator variables in neurofeedback training in children with ADHD. EUNTEHYDIS international conference on ADHD*, Stockholm, Sweden.

Mar. 2015 **Zuberer, A.**Brandeis, D., Drechsler, R., *Knowing and changing mental states: New insights from a selfappraisal task of spontaneous arousal*, Neurizons 2015: Notes of an Orchestrated Brain. Max Planck Institute for Biophysical chemistry, Göttingen, Germany.

Oct. 2014 Minder, F. **Zuberer, A.,** Wirth, L., Drechsler, R., *Standardized school observations and teacher ratings of children with Attention-Deficit*, Hyperactivity Disorder. DGKJP congress, Munich, Germany.